ANALYSIS OF VARIANCE

FOR DATA HAVING

UNEQUAL NUMBERS OF OBSERVATIONS

IN THE SUBCLASSES

Notes for a short course based on "Linear Models"

by

S. R. Searle
Biometrics Unit
Cornell University, Ithaca, New York

Paper No. BU-658a-M in the Biometrics Unit, being BU-658-M with supplemental pages 52a, 58a,b, 61a,b and 62a.

TABLE OF CONTENTS

CHAPTER 1

1. INTRODUCTION

These are notes for a short course, based on material in the text "Linear Models", by Searle, Wiley, 1971. The notes are just that, notes: and they are not to be viewed as anything else. They are to be used in conjunction with the text, which is referenced frequently, in the form LM 262, for example, meaning page 262.

1.1. Analysis of Variance

The object of statistical analysis of data is to elicit information from data, especially to sort out information from random variation; i.e., to distinguish signal from noise. Analysis of variance is one of the oldest techniques for doing this. It had its origin in the analysis of agricultural experiments, its early development being the inspiration of R. A. Fisher at Rothamsted Experiment Station during the 1920's and 1930's.

The basic idea of analysis of variance is that of partitioning the sum of squares of a set of observations into sources of variation that are meaningful, as well as a source that cannot be explained. (The former constitute the signals and the latter is the noise.) Thus analysis of variance might well be called partitioning of variation. Sometimes the partitioning is simple, both algebraically and computationally, and sometimes it is quite difficult. This course deals with difficult cases, although four easy cases are reviewed first, as a starting point.

1.2. Hypothetical Data as Illustrations

There are at least two ways of utilizing data in the process of learning statistical methods. One is to have a variety of sets of real-life data and to use them for illustrating the different statistical methods encountered; i.e., case-studies. This is the procedure adopted in most texts on methodology, e.g., Johnson and Leone, *Statistics and Experimental Design in Engineering and the Physical Sciences*, Wiley, 1964. It is ideal for providing motivation, for giving a real feel for what statistical methods do, and for learning how to interpret analyses. In contrast, once one has a good understanding of basic concepts (e.g., sums of squares, analysis of variance tables, normality, chi-square distributions, independence, F-statistics, hypothesis testing and so on), the learning of complicated applications of them may be enhanced by using not real-life data, but hypothetical data. With a motivation to learn and an understanding of basics already in place, there is no need for having real-life data just to provide these ingredients of the learning process. Furthermore, for complicated analyses, where learning just what the correct calculations are (and also learning about similar-looking and not quite correct ones), calculations using real-life data have the disadvantage of being almost as difficult to follow as is the algebra they are intended to illustrate. On the other hand, hypothetical data can be constructed so that some of the calculations are easy to do, and hence easy to follow, and so their use as illustration of the algebra is instructive.

Once the calculation techniques of new statistical methods have been learned, and reinforced by illustration with hypothetical data, the methods can be applied to real-life data and, from one's earlier knowledge of concepts, interpretation can be appropriately made. This is the approach taken in this course and in the text. It is assumed that basic concepts of analysis of variance are known and understood. Illustrations of techniques are therefore all in terms of hypothetical data.

It has been said that one does not learn how to solve quadratic equations by first considering something like $264.38162x^2 - 19.71043x - .00871379 = 0$. Even if that is an equation with real-life application, and $x^2 - 7x + 12 = 0$ is not, the latter is certainly a better vehicle than the former for first learning about quadratic equations. That is the attitude taken here.

1.3. Vocabulary

Example: Suppose 3 brands of loom are used in a textile plant. During 2 randomly chosen hours one day, the number of minutes of machine break-down is recorded for each brand of loom. The number of observations is therefore 2 per loom, as summarized in Table 1.

Table 1

Number of Observations

Brand of Loom		
A	B	C
2	2	2

If the same procedure were carried out in 5 different textile plants the numbers of observations would be 2 per brand of loom in each plant, as shown in Table 2.

Table 2

Number of Observations

Textile Plant	Brand of Loom A	B	C
1	2	2	2
2	2	2	2
3	2	2	2
4	2	2	2
5	2	2	2

a. Factors

The categories by which data can be classified are called factors. In Table 1, there is one factor: brand of loom. In Table 2, there are two factors: textile plant and brand of loom.

b. Levels of a factor

The sub-categories of each factor are called levels. In Tables 1 and 2, there are 3 levels of the factor brand of loom, and in Table 2 there are 5 levels of the factor textile plant.

c. Subclasses

The subclasses of a set of data are the sub-categories defined by the intersections of the levels of the factors. In Table 1 the subclasses are the levels of the factor brand of loom. In Table 2 the subclasses are the 15 sub-categories defined by the 3 levels of brand and the 5 levels of textile plant; e.g., plant 1, brand A is one subclass; plant 3, brand B is another, and so on.

d. Balanced data

We define balanced data as data in which all the subclasses have the same number of observations; i.e., equal-subclass-numbers data. Tables 1 and 2 are examples.

Caution: The word "balance" is to be found in several other contexts in describing data; e.g., balanced incomplete blocks, variance balance, and so on.

Suppose in the example of Tables 1 and 2 that the textile plants operated 24 hours a day, and instead of recording machine stoppages for just two randomly chosen hours during a day that records were taken for different numbers of randomly chosen hours throughout a 24-hour day, different for each brand in each plant. The numbers of observations might be those shown in Tables 3 and 4.

Table 3

Numbers of Observations

Brand of Loom		
A	B	C
6	4	5

Table 4

Numbers of Observations*

Textile Plant	Brand of Loom A	B	C	Total
1	6	4	5	15
2	-	3	2	5
3	5	-	7	12
4	-	-	9	9
5	3	4	-	7
Total	14	11	23	48

* A dash represents no observation.

e. Unbalanced data

Unbalanced data are those having differing numbers of observations in the subclasses; i.e., unequal-subclass-numbers data. Some subclasses may be empty, and contain no data at all. Tables 3 and 4 are examples.

f. Designed experiments

One source of good data is a well designed and well executed experiment. Methods of designing good experiments are not part of this course; it is concerned with analysis of data once they are available. The source of one's data is assumed to be known and the data are assumed to contain information about the problem at hand. (In practice, no statistician worth his salt should make this assumption about data: its source should always be well understood. Valid interpretation of data demands that both its source and its analysis be satisfactory. In order to concentrate on analysis, this course assumes the source of data is satisfactory.)

g. Planned unbalancedness

A feature of many experimental designs is that the data they yield are unbalanced, but with the unbalancedness being carefully planned as part of the design so as to leave the analysis of variance easy to calculate and interpret without serious loss of information because of it. Balanced incomplete block experiments are of this nature. An example is as follows:

Table 5

Number of Observations

(Balanced incomplete block data)

Textile Plant	Brand of Loom				Total
	A	B	C	D	
1	1	1	-	-	2
2	1	-	1	-	2
3	1	-	-	1	2
4	-	1	1	-	2
5	-	1	-	1	2
6	-	-	1	1	2
Total	3	3	3	3	12

Data from the pattern of observations shown in Table 5 could be analyzed by the methods for unbalanced data that will be presented in these notes. But the well-ordered pattern of the unbalancedness will reduce those calculations from their usual complexity for any sort of unbalancedness to the familiar and tractable form that is well known for balanced incomplete block experiments.

h. Missing data

In the execution of well designed experiments, intended observations are sometimes lost or never obtained; e.g., laboratory animals die, and machines break down. The resulting data then have unequal numbers of observations in the subclasses. But if only a very few observations are missing from these otherwise balanced data, techniques known as missing cell techniques can be used to make the data balanced for purposes of analysis. (See LM 362.)

i. Descriptive data

In juxtaposition to data available from experiments planned in advance, are data that are simply there, data that are descriptive of certain situations, data that are available to be recorded, analyzed and interpreted if so desired. Examples are legion: hospital data, population data, clinical trial data, medical and epidemiological data, meteorological and astronomy data, shipping and transport data, production and construction data, and so on and so on. In all cases such data can usually be classified according to several factors, each with a number of levels. But by their very nature, such data will rarely (if ever) fall into these subclasses in equal numbers. Furthermore, by their very definition, some subclasses are bound to be empty; for example, the subclass of a married teenager with 5 children, on welfare, owning a $400,000 mortgage-free house in Georgetown is unlikely to contain observations.

Descriptive data of this nature are therefore often very unbalanced. They are also usually quite voluminous and consequently are best handled using computer technology, both for storage and analysis.

j. Unbalanced data and analysis of variance

Armed with a knowledge of analysis of variance techniques for balanced data, one might wonder if the apparently small change in the characteristic of being unbalanced rather than balanced makes much difference to the algebra, arithmetic, and interpretation of the analysis of variance technique. Perhaps surprisingly, it does make a difference — a very big difference. Plainly put, the situation is as follows:

Data	Analysis of variance
Balanced data	Easy
Unbalanced data	Difficult

k. Origins of analysis of variance: analyzing means

In the example of Table 2, the case of 2 observations on the number of breakdowns per hour on each of 3 brands of loom in 5 textile plants, let

$$y_{ijk} = \text{k'th observation in plant i on brand j}$$

for $i = 1, 2, \cdots, 5$, $j = 1, 2, 3$ and $k = 1, 2$. The total sum of squares of the y_{ijk}'s about their mean

$$\bar{y}_{...} = \sum_{i=1}^{5} \sum_{j=1}^{3} \sum_{k=1}^{2} y_{ijk}/30 = \text{grand mean}$$

is $SST_m = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{...})^2$. Analysis of variance calculations consist of partitioning this in terms of the other means available from the data:

$$\bar{y}_{i..} = \sum_{j=1}^{3} \sum_{k=1}^{2} y_{ijk}/6 \quad = \text{mean for plant } i$$

$$\bar{y}_{.j.} = \sum_{i=1}^{5} \sum_{k=1}^{2} y_{ijk}/10 = \text{mean for brand } j$$

$$\bar{y}_{ij.} = \sum_{k=1}^{2} y_{ijk}/2 \quad = \text{mean for plant } i, \text{ brand } j \text{ .}$$

The basis of the partitioning is the <u>mathematical identity</u>

$$\sum_{i}\sum_{j}\sum_{k}(y_{ijk} - \bar{y}_{...})^2$$

$$\equiv \sum_{i}\sum_{j}\sum_{k}(\bar{y}_{i..} - \bar{y}_{...})^2 + \sum_{i}\sum_{j}\sum_{k}(\bar{y}_{.j.} - \bar{y}_{...})^2$$

$$+ \sum_{i}\sum_{j}\sum_{k}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2 + \sum_{i}\sum_{j}\sum_{k}(y_{ijk} - \bar{y}_{ij.})^2,$$

an identity which is straight algebra and involves no statistics whatever. The Analysis of Variance table summarizes this identity and identifies each sum of squares with a name indicating the source of variation it is measuring. Its familiar form is shown in Table 6.

Table 6

Analysis of Variance of
data based on Table 2

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F
Plant	4	$SSA = \sum\sum\sum_{ijk}(\bar{y}_{i..} - \bar{y}_{...})^2$	MSA = SSA/4	MSA/MSE
Brand	2	$SSB = \sum\sum\sum_{ijk}(\bar{y}_{.j.} - \bar{y}_{...})^2$	MSB = SSB/2	MSB/MSE
Interaction	8	$SSAB = \sum\sum\sum_{ijk}(\bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...})^2$	MSAB = SSAB/8	MSAB/MSE
Residual	15	$SSE = \sum\sum\sum_{ijk}(y_{ijk} - \bar{y}_{ij.})^2$	MSE = SSE/15	
Total	29	$SST_m = \sum\sum\sum_{ijk}(y_{ijk} - \bar{y}_{...})^2$		

So far as the sums of squares are concerned, the table represents no more than a summary of the partitioning of SST_m (total sum of squares about the mean) into the four other sums of squares of the identity.

To this summary, Fisher applied the distributional property of normality, assuming the y_{ijk}'s to be normally and independently distributed with equal variance. Based on this assumption he showed that each of the sums of squares SSA, SSB, SSAB and SSE has a distribution that is proportional to a χ^2-distribution; and these sums of squares are distributed independently of one another. This leads to ratios of the mean squares having F-distributions, non-central F's, which under appropriate hypotheses reduce to central F-distributions. Comparing the calculated F's with tabulated F's therefore provides tests of those hypotheses.

The analysis of variance table is essentially just a convenient summary of calculations. It is the assumption of normality and the concept of hypothesis testing that leads to being able to apply that concept using the calculated F's. Furthermore, all of this is just in terms of sums of squares of deviations of means from each other.

<u>*l*. Linear models</u>

It seems apparent that Fisher's work was based on deviations of means as just illustrated, and that he did not use linear models as has been customary in the teaching of statistics during the last twenty years or so. An example of such a model, for the illustration of Table 2 is

$$y_{ijk} = \mu + p_i + b_j + (pb)_{ij} + e_{ijk}$$

where μ is a general mean, p_i is the effect due to the i'th plant, b_j is the effect due to the j'th brand, $(pb)_{ij}$ is the interaction effect and e_{ijk} is the error term. Although models of this nature have done much to clarify our understanding of analysis of variance techniques, they have also, as we shall see, muddied the waters in certain situations.

The prime topic of this course is the analysis of variance of unbalanced data. As prelude thereto we give, in note form, a brief summary of the analysis of variance of three simple cases of balanced data. Unbalanced data for the same three cases are considered subsequently.

CHAPTER 2

2. FOUR BASIC BALANCED DATA CASES

To introduce notation and to provide a brief review of analysis of variance of balanced data we here provide a short summary of the following four models:

2.1. The 1-way classification.

2.2. The 2-way nested classification.

2.3. The 2-way crossed classification with $n = 1$ observation per cell (no interaction).

2.4. The 2-way crossed classification with $n > 1$ observation per cell (with interaction).

In designed experiments an example of the first of these is the completely randomized design, and examples of the last two are randomized complete block designs.

All four models are summarized under the following headings:

a. Example

b. Model

c. Model equations

d. Normal equations

e. Restrictions on model

f. Constraints on solution

g. Solution

h. Analysis of variance

i. Estimated residual variance

j. Hypothesis testing using F's

k. Orthogonal contrasts

ℓ. Estimated differences between levels

m. Variance of an estimated difference between levels.

2.1. The 1-way classification: balanced data. (Completely randomized design)

a. Example: 2 observations on each of 3 brands (Table 1).

Table 7

A	B	C	
6	9	6	
10	5	18	
$y_{1.} = 16$	$y_{2.} = 14$	$y_{3.} = 24$	$y_{..} = 54$
$\bar{y}_{1.} = 8$	$\bar{y}_{2.} = 7$	$\bar{y}_{3.} = 12$	$\bar{y}_{..} = 9$

b. Model: The usual linear model is

$$y_{ij} = \mu + \alpha_i + e_{ij}$$

where μ is a general mean, α_i is the effect due to the i'th brand and e_{ij} is a random error term assumed to have zero mean and variance σ^2. In general, $i = 1, \cdots, a$ and $j = 1, \cdots, n$ and for our example $a = 3$ and $n = 2$.

c. Model equations: The model equations for the data are:

$$\begin{aligned}
y_{11} &= 6 = \mu + \alpha_1 &&&& + e_{11} \\
y_{12} &= 10 = \mu + \alpha_1 &&&& + e_{12} \\
y_{21} &= 9 = \mu && + \alpha_2 && + e_{21} \\
y_{22} &= 5 = \mu && + \alpha_2 && + e_{22} \\
y_{31} &= 6 = \mu &&& + \alpha_3 & + e_{31} \\
y_{32} &= 18 = \mu &&& + \alpha_3 & + e_{32}
\end{aligned}$$

which, written in matrix form

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$$

are:

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 6 \\ 10 \\ 9 \\ 5 \\ 6 \\ 18 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \end{bmatrix}$$

d. Normal equations:

$$\underset{\sim}{X}'\underset{\sim}{X}\hat{\underset{\sim}{b}} = \underset{\sim}{X}'\underset{\sim}{y}$$

$$\begin{aligned} 6\hat{\mu} + 2\hat{\alpha}_1 + 2\hat{\alpha}_2 + 2\hat{\alpha}_3 &= y_{..} = 54 \\ 2\hat{\mu} + 2\hat{\alpha}_1 \qquad\qquad\qquad &= y_{1.} = 16 \\ 2\hat{\mu} \qquad\quad + 2\hat{\alpha}_2 \qquad\quad &= y_{2.} = 14 \\ 2\hat{\mu} \qquad\qquad\qquad + 2\hat{\alpha}_3 &= y_{3.} = 24 \end{aligned}$$

$$\begin{bmatrix} 6 & 2 & 2 & 2 \\ 2 & 2 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \end{bmatrix} = \begin{bmatrix} y_{..} \\ y_{1.} \\ y_{2.} \\ y_{3.} \end{bmatrix} = \begin{bmatrix} 54 \\ 16 \\ 14 \\ 24 \end{bmatrix}$$

Note: Last 3 equations add to first.
Therefore there is no unique solution, but infinitely many solutions.
Obtain a solution using a restriction.

e. Restrictions on model: $\alpha_1 + \alpha_2 + \alpha_3 = 0$

f. Constraints on solution: $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = 0$

g. Solution:

$$\hat{\underset{\sim}{b}} = \begin{bmatrix} \hat{\mu} \\ \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \hat{\alpha}_3 \end{bmatrix} = \begin{bmatrix} \bar{y}_{..} \\ \bar{y}_{1.} - \bar{y}_{..} \\ \bar{y}_{2.} - \bar{y}_{..} \\ \bar{y}_{3.} - \bar{y}_{..} \end{bmatrix} = \begin{bmatrix} 9 \\ 8 - 9 \\ 7 - 9 \\ 12 - 9 \end{bmatrix} = \begin{bmatrix} 9 \\ -1 \\ -2 \\ 3 \end{bmatrix}$$

Note: $\hat{\alpha}_1 + \hat{\alpha}_2 + \hat{\alpha}_3 = -1 - 2 + 3 = 0.$

h. Analysis of variance:

Table 8, page 17.

i. Estimated residual variance:

$$\hat{\sigma}^2 = \text{MSE} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{n} (y_{ij} - \bar{y}_{i.})^2}{a(n-1)} = 88/3 .$$

j. Hypothesis testing using F's:

F(A) tests $H: \alpha_1 = \alpha_2 = \alpha_3$

F(M) tests $H: E(\bar{y}_{..}) = 0$.

This is equivalent to $H: \mu + (\alpha_1 + \alpha_2 + \alpha_3)/3 = 0.$

But the model includes the restriction $\alpha_1 + \alpha_2 + \alpha_3 = 0.$

Therefore this Hypothesis is $H: \mu = 0.$

Note that $F(M) = t^2$ where t is the t-statistic for testing the same hypothesis. This is so because

$$t = \frac{\bar{y}_{..}}{\sqrt{\hat{v}(\bar{y})}} = \frac{\bar{y}_{..}}{\sqrt{\hat{\sigma}^2/an}} = \sqrt{\frac{an\bar{y}_{..}^2}{\hat{\sigma}^2}} = \sqrt{\frac{an\bar{y}_{..}^2}{\text{MSE}}} = \sqrt{F(M)} .$$

Table 8

Analysis of Variance of Data from Table 7

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	F-Statistics
Traditional form				
Machines	$a-1 = 2$	$SSA = \sum_{ij}\sum(\bar{y}_{i.} - \bar{y}_{..})^2 = 28$	$MSA = 28/2$	$F(A) = \frac{28}{2}/\frac{88}{3}$
Residual	$a(n-1) = 3$	$SSE = \sum_{ij}\sum(\bar{y}_{ij} - \bar{y}_{i.})^2 = 88$	$MSE = 88/3$	
Total (c.f.m.)	$an-1 = 5$	$SST_m = \sum_{ij}\sum(\bar{y}_{ij} - \bar{y}_{..})^2 = 116$		
A form that includes the mean				
Mean	$1 = 1$	$SSM = an\bar{y}^2_{..} = 486$	$MSM = 486$	$F(M) = 486/\frac{88}{3}$
Machines	$a-1 = 2$	$SSA =$ as above $= 28$	$MSA = 28/2$	$F(A) = \frac{28}{2}/\frac{88}{3}$
Residual	$a(n-1) = 3$	$SSE =$ as above $= 88$	$MSE = 88/3$	
Total	$an = 6$	$SST = \sum_{ij}\sum y^2_{ij} = 602$		
A form just for fitting the whole model $y_{ij} = \mu + \alpha_i + e_{ij}$				
Model	$a = 3$	$SSR = \underset{\sim}{\hat{b}}'\underset{\sim}{X}'\underset{\sim}{y} = 514$	$MSR = 514/3$	$F(R) = \frac{514}{3}/\frac{88}{3}$
Residual	$an-a = 3$	$SSE = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{\hat{b}}'\underset{\sim}{X}'\underset{\sim}{y} = 88$	$MSE = 88/3$	
Total	$an = 6$	$SST = \underset{\sim}{y}'\underset{\sim}{y} = 602$		

k. Orthogonal contrasts: SSA can be partitioned into sums of squares for orthogonal 1-degree-of-freedom contrasts.

Examples

$$\begin{array}{lll}
\alpha_1 - \alpha_2 & 2[(8 - 7)^2/(1 + 1)] & = 1 \\
\alpha_1 + \alpha_2 - 2\alpha_3 & 2[(8 + 7 - 24)^2/(1 + 1 + 4)] & = \underline{27} \\
 & & \underline{28} = \text{SSA}
\end{array}$$

or

$$\begin{array}{lll}
\alpha_1 - \alpha_3 & 2[(8 - 12)^2/(1 + 1)] & = 16 \\
\alpha_1 + \alpha_3 - 2\alpha_2 & 2[(8 + 12 - 14)^2/(1 + 1 + 4)] & = \underline{12} \\
 & & \underline{28} = \text{SSA}
\end{array}$$

ℓ. Estimated differences between levels:

$$\begin{array}{ll}
\text{A vs. B:} & \hat{\alpha}_1 - \hat{\alpha}_2 = \bar{y}_{1.} - \bar{y}_{2.} = 8 - 7 = 1 \\
\text{A vs. C:} & \hat{\alpha}_1 - \hat{\alpha}_2 = \bar{y}_{1.} - \bar{y}_{3.} = 8 - 12 = -4 \\
\text{B vs. C:} & \hat{\alpha}_2 - \hat{\alpha}_3 = \bar{y}_{2.} - \bar{y}_{3.} = 7 - 12 = -5
\end{array}$$

m. Variance of an estimated difference between levels:

$$v(\bar{y}_{i.} - \bar{y}_{i'.}) = \sigma^2(\tfrac{1}{4} + \tfrac{1}{4}) = \tfrac{1}{2}\sigma^2 \ .$$

These permit confidence intervals and multiple range tests to be computed in the usual way.

2.2. The 2-way nested (hierarchical) classification: balanced data.

a. Example: Suppose productivity of assembly-line workers is to be studied in 2 different assembly lines within each major plant of the 3 major automobile manufacturers. At each line, 2 workers are chosen at random. Let

$$y_{ijk} = \text{productivity of worker k in line j in plant i.}$$

In general we may have a manufacturers, b lines in each and n workers from each line, so that $i = 1, 2, \cdots, a$, $j = 1, 2, \cdots, b$ and $k = 1, 2, \cdots, n$. In our case $a = 3$, $b = 2$ and $n = 2$.

Suppose the data are as shown in Table 9.

Table 9

Data for a 2-way nested classification

	Plant					
	A		B		C	
	Line		Line		Line	
	1	2	1	2	1	2
	3	6	3	4	2	9
	7	8	5	8	4	1
Line totals:	$y_{11.} = 10$	$y_{12.} = 14$	$y_{21.} = 8$	$y_{22.} = 12$	$y_{31.} = 6$	$y_{32.} = 10$
Plant totals:	$y_{1..} = 24$		$y_{2..} = 20$		$y_{3..} = 16$	
Line means:	$\bar{y}_{11.} = 5$	$\bar{y}_{12.} = 7$	$\bar{y}_{21.} = 4$	$\bar{y}_{22.} = 6$	$\bar{y}_{31.} = 3$	$\bar{y}_{32.} = 5$
Plant means:	$\bar{y}_{1..} = 6$		$\bar{y}_{2..} = 5$		$\bar{y}_{3..} = 4$	

Grand total: $y_{...} = 60$

Overall mean: $\bar{y}_{...} = 5$

b. Model: The customary linear model is

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ijk}$$

where μ is a general mean, α_i is the effect due to manufacturer i, β_{ij} is the effect due to line j within manufacturer i, and e_{ijk} is the usual error term, having variance σ^2.

c. Model equations: In matrix form, $\underset{\sim}{y} = \underset{\sim\sim}{Xb} + \underset{\sim}{e}$, the model equations are

$$\underset{\sim}{y} = \begin{bmatrix} 3 \\ 7 \\ 6 \\ 8 \\ 3 \\ 5 \\ 4 \\ 8 \\ 2 \\ 4 \\ 9 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 & . & . & 1 & . & . & . & . & . \\ 1 & 1 & . & . & 1 & . & . & . & . & . \\ 1 & 1 & . & . & . & 1 & . & . & . & . \\ 1 & 1 & . & . & . & 1 & . & . & . & . \\ 1 & . & 1 & . & . & . & 1 & . & . & . \\ 1 & . & 1 & . & . & . & 1 & . & . & . \\ 1 & . & 1 & . & . & . & . & 1 & . & . \\ 1 & . & 1 & . & . & . & . & 1 & . & . \\ 1 & . & . & 1 & . & . & . & . & 1 & . \\ 1 & . & . & 1 & . & . & . & . & 1 & . \\ 1 & . & . & 1 & . & . & . & . & . & 1 \\ 1 & . & . & 1 & . & . & . & . & . & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \\ \beta_{31} \\ \beta_{32} \end{bmatrix} + \underset{\sim}{e} \,.$$

Dots in a matrix represent zeros.

d. Normal equations: The normal equations $\underset{\sim}{X}'\underset{\sim}{X}\hat{\underset{\sim}{b}} = \underset{\sim}{X}'\underset{\sim}{y}$ for these data are as follows:

$$
\begin{array}{llllllllllllll}
12\hat{\mu} & + 4(\hat{\alpha}_1 & + \hat{\alpha}_2 & + \hat{\alpha}_3) & + 2(\hat{\beta}_{11} & + \hat{\beta}_{12} & + \hat{\beta}_{21} & + \hat{\beta}_{22} & + \hat{\beta}_{31} & + \hat{\beta}_{32}) & = y_{\cdot\cdot\cdot} & = 60 \\
4\hat{\mu} & + 4\hat{\alpha}_1 & & & + 2(\hat{\beta}_{11} & + \hat{\beta}_{12}) & & & & & = y_{1\cdot\cdot} & = 24 \\
4\hat{\mu} & & + 4\hat{\alpha}_2 & & & & + 2(\hat{\beta}_{21} & + \hat{\beta}_{22}) & & & = y_{2\cdot\cdot} & = 20 \\
4\hat{\mu} & & & + 4\hat{\alpha}_3 & & & & & + 2(\hat{\beta}_{31} & + \hat{\beta}_{32}) & = y_{3\cdot\cdot} & = 16 \\
2\hat{\mu} & + 2\hat{\alpha}_1 & & & + 2\hat{\beta}_{11} & & & & & & = y_{11\cdot} & = 10 \\
2\hat{\mu} & + 2\hat{\alpha}_1 & & & & + 2\hat{\beta}_{12} & & & & & = y_{12\cdot} & = 14 \\
2\hat{\mu} & & + 2\hat{\alpha}_2 & & & & + 2\hat{\beta}_{21} & & & & = y_{21\cdot} & = 8 \\
2\hat{\mu} & & + 2\hat{\alpha}_2 & & & & & + 2\hat{\beta}_{22} & & & = y_{22\cdot} & = 12 \\
2\hat{\mu} & & & + 2\hat{\alpha}_3 & & & & & + 2\hat{\beta}_{31} & & = y_{31\cdot} & = 6 \\
2\hat{\mu} & & & + 2\hat{\alpha}_3 & & & & & & + 2\hat{\beta}_{32} & = y_{32\cdot} & = 10
\end{array}
$$

The general form of these equations is

$$abn\hat{\mu} + bn\sum_{i=1}^{a}\hat{\alpha}_i + n\sum_{i=1}^{a}\sum_{j=1}^{b}\hat{\beta}_{ij} = y_{\cdot\cdot\cdot}$$

$$bn\hat{\mu} + bn\hat{\alpha}_i + n\sum_{j=1}^{b}\hat{\beta}_{ij} = y_{i\cdot\cdot}\,,$$

one equation for each $i = 1, \cdots, a$

$$n\hat{\mu} + n\hat{\alpha}_i + n\hat{\beta}_{ij} = y_{ij\cdot}\,,$$

one equation for each combination of $i = 1, \cdots, a$ and $j = 1, \cdots, b$.

e. Restrictions on model: It is customary in this model, with balanced data as we have here, to include the following restrictions on the parameters as part of the model:

$$\sum_{i=1}^{a}\alpha_i = 0 \quad \text{and} \quad \sum_{j=1}^{b}\beta_{ij} = 0 \quad \text{for all } i = 1, \cdots, a.$$

f. Constraints on solution: Solutions to the normal equations are obtained using similar limitations on the solutions:

$$\sum_{i=1}^{a} \hat{\alpha}_i = 0 \quad \text{and} \quad \sum_{j=1}^{b} \hat{\beta}_{ij} = 0 \quad \text{for all } i = 1, \cdots, a.$$

g. Solution: The resulting solution vector $\hat{\underset{\sim}{b}}$ has as its elements

$$\hat{\mu} = \bar{y}_{...}, \quad \hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...} \quad \text{and} \quad \hat{\beta}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} .$$

For our example this is

$$\hat{\mu} = 5 \qquad \begin{aligned} \hat{\alpha}_1 &= 6 - 5 = 1 \\ \hat{\alpha}_2 &= 5 - 5 = 0 \\ \hat{\alpha}_3 &= 4 - 5 = -1 \end{aligned} \qquad \begin{aligned} \hat{\beta}_{11} &= 5 - 6 = -1 \\ \hat{\beta}_{12} &= 7 - 6 = 1 \\ \hat{\beta}_{21} &= 4 - 5 = -1 \\ \hat{\beta}_{22} &= 6 - 5 = 1 \\ \hat{\beta}_{31} &= 3 - 4 = -1 \\ \hat{\beta}_{32} &= 5 - 4 = 1 \end{aligned}$$

h. Analysis of variance: Table 10 shows both the general analysis of variance and that for the data of the example. (See Table 10, page 23.)

i. Estimated residual variance: The residual error variance σ^2 is estimated by the error mean square

$$\hat{\sigma}^2 = \frac{\sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{ij.})^2}{ab(n-1)} = \frac{54}{6} = 9 .$$

j. Hypothesis testing using F's: In the presence of the restrictions included in the model, the F-statistics in the analysis of variance table can be used for testing hypotheses as follows:

F(M) tests $H : \mu = 0$, i.e., $H : E(\bar{y}_{...}) = 0$, using the restrictions

F(A) tests $H : \alpha_i$'s all equal

F(B:A) tests H : equality of β_{ij}'s within each α_i .

Table 10

Analysis of Variance for a 2-way Nested Classification, Balanced Data.

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Square	F-Statistics
General case				
Mean	1	$SSM = abn\bar{y}^2_{...}$	$MSM = SSM$	$F(M) = \frac{MSM}{MSE}$
Plant	a-1	$SSA = bn \sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2$	$MSA = \frac{MSA}{a-1}$	$F(A) = \frac{MSA}{MSE}$
Lines within Plant	a(b-1)	$SSB{:}A = n \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{ij.} - \bar{y}_{i..})^2$	$MSB{:}A = \frac{SSB{:}A}{a(b-1)}$	$F(B{:}A) = \frac{MSB{:}A}{MSE}$
Residual	ab(n-1)	$SSE = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} (y_{ijk} - \bar{y}_{ij.})^2$	$MSE = \frac{SSE}{ab(n-1)}$	
Total	abn	$SST = \sum_{i=1}^{a} \sum_{j=1}^{b} \sum_{k=1}^{n} y^2_{ijk}$		
Example (Table 9)				
Mean	1	SSM = 300	MSM = 300	F(M) = 10/3
Plant	2	SSA = 8	MSA = 4	F(A) = 4/9
Lines within Makes	3	SSB:A = 12	MSB:A = 4	F(B:A) = 4/9
Residual	6	SSE = 54	MSE = 9	
Total	12	SST = 374		

k. Orthogonal contrasts: Sums of squares SSA and SSB:A can be partitioned into sums of squares for orthogonal 1-degree-of-freedom contrasts.

Examples

Partitioning SSA

$\alpha_1 - \alpha_2$	$4(6-5)^2/(1+1)$	$= 2$
$\alpha_1 + \alpha_2 - 2\alpha_3$	$4(6+5-8)^2/(1+1+4)$	$= 6$
		$8 = \text{SSA}$

Partitioning SSB:A

$\beta_{11} - \beta_{12}$	$2(5-7)^2/(1+1)$	$= 4$
$\beta_{21} - \beta_{22}$	$2(4-6)^2/(1+1)$	$= 4$
$\beta_{31} - \beta_{32}$	$2(3-5)^2/(1+1)$	$= 4$
		$12 = \text{SSB:A}$

ℓ. Estimated differences between levels:

$$\hat{\alpha}_i - \hat{\alpha}_{i'} = \bar{y}_{i..} - \bar{y}_{i'..} \quad \text{for } i \neq i'$$

$$\hat{\beta}_{ij} - \hat{\beta}_{ij'} = \bar{y}_{ij.} - \bar{y}_{ij'.} \quad \text{for any } i, \text{ and } j \neq j' .$$

m. Variance of an estimated difference between levels:

$$v(\hat{\alpha}_i - \hat{\alpha}_{i'}) = 2\sigma^2/bn .$$

$$v(\hat{\beta}_{ij} - \hat{\beta}_{ij'}) = 2\sigma^2/n .$$

2.3. The 2-way crossed classification with 1 observation per cell.

a. Example: Suppose 3 brands of loom are each tested just once in 2 plants with

$$y_{ij} = \text{observation in plant } i \text{ on brand } j.$$

In general we may have a plants and b brands so that $i = 1, \cdots, a$ and $j = 1, \cdots, b$. In our case $a = 2$ and $b = 3$.

We analyze the following data.

Table 11

Data for a 2-way crossed classification

Textile Plant	Brand A	B	C	Total	Mean
1	11	9	7	27	9
2	9	13	11	33	11
Total	20	22	18	60	
Mean	10	11	9		10

b. Model: The traditional linear model for this kind of a situation is

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij}$$

where μ is a general mean, α_i is the effect due to plant i, β_j is the effect due to brand j and e_{ij} is the usual random error term assumed to have variance σ^2.

c. Model equations: The matrix form $\underset{\sim}{y} = \underset{\sim\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$ of the normal equations is

$$\underset{\sim}{y} = \begin{bmatrix} 11 \\ 9 \\ 7 \\ 9 \\ 13 \\ 11 \end{bmatrix} = \begin{bmatrix} 1 & 1 & . & 1 & . & . \\ 1 & 1 & . & . & 1 & . \\ 1 & 1 & . & . & . & 1 \\ 1 & . & 1 & 1 & . & . \\ 1 & . & 1 & . & 1 & . \\ 1 & . & 1 & . & . & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \beta_1 \\ \beta_2 \\ \beta_3 \end{bmatrix} + \underset{\sim}{e} .$$

d. Normal equations: The normal equations $\underset{\sim}{X}'\underset{\sim}{X}\hat{\underset{\sim}{b}} = \underset{\sim}{X}'\underset{\sim}{y}$ are

$$\begin{aligned}
6\hat{\mu} + 3\hat{\alpha}_1 + 3\hat{\alpha}_2 + 2\hat{\beta}_1 + 2\hat{\beta}_2 + 2\hat{\beta}_3 &= y_{..} = 60 \\
3\hat{\mu} + 3\hat{\alpha}_1 \phantom{+ 3\hat{\alpha}_2} + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 &= y_{1.} = 27 \\
3\hat{\mu} \phantom{+ 3\hat{\alpha}_1} + 3\hat{\alpha}_2 + \hat{\beta}_1 + \hat{\beta}_2 + \hat{\beta}_3 &= y_{2.} = 33 \\
2\hat{\mu} + \hat{\alpha}_1 + \hat{\alpha}_2 + 2\hat{\beta}_1 &= y_{.1} = 20 \\
2\hat{\mu} + \hat{\alpha}_1 + \hat{\alpha}_2 + 2\hat{\beta}_2 &= y_{.2} = 22 \\
2\hat{\mu} + \hat{\alpha}_1 + \hat{\alpha}_2 + 2\hat{\beta}_3 &= y_{.3} = 18
\end{aligned}$$

The general form of these is

$$ab\hat{\mu} + b\sum_{i=1}^{a}\hat{\alpha}_i + a\sum_{j=1}^{b}\hat{\beta}_j = y_{..}$$

$$b\hat{\mu} + b\hat{\alpha}_i + \Sigma\hat{\beta}_j = y_{i.}\,,$$

one equation for each $i = 1, \cdots, a$

$$a\hat{\mu} + \sum_{i=1}^{a}\hat{\alpha}_i + b\hat{\beta}_j = y_{.j}\,,$$

one equation for each $j = 1, \cdots, b$.

e. Restrictions on model: With balanced data, as here, the customary restrictions used as part of the model are

$$\sum_{i=1}^{a}\alpha_i = 0 \quad \text{and} \quad \sum_{j=1}^{b}\beta_j = 0\,.$$

f. Constraints on solution: Solutions to the normal equations are obtained with the use of constraints on the solutions similar to the restrictions on the model:

$$\sum_{i=1}^{a}\hat{\alpha}_i = 0 \quad \text{and} \quad \sum_{j=1}^{b}\hat{\beta}_j = 0\,.$$

g. Solution: Elements of $\hat{\underset{\sim}{b}}$ are then

$$\hat{\mu} = \bar{y}_{..}, \quad \hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} \quad \text{and} \quad \hat{\beta}_j = \bar{y}_{.j} - \bar{y}_{..} .$$

For the example of Table 11

$$\hat{\mu} = 10 \qquad \hat{\alpha}_1 = 9 - 10 = -1 \qquad \hat{\beta}_1 = 10 - 10 = 0$$

$$\hat{\alpha}_2 = 11 - 10 = 1 \qquad \hat{\beta}_2 = 11 - 10 = 1$$

$$\hat{\beta}_3 = 9 - 10 = -1 .$$

h. Analysis of variance: Table 12 shows the analysis of variance for the general case and for the illustrative data of Table 11. (See Table 12, page 28.)

i. Estimated residual variance:

$$\hat{\sigma}^2 = \text{MSE} = \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} (y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2}{(a - 1)(b - 1)} = 6 .$$

j. Hypothesis testing using F's: Under the restrictions included as part of the model, the F-statistics in the analysis of variance table can be used to test hypotheses as follows:

F(M) tests $H : \mu = 0$, i.e., $H : E(\bar{y}_{..}) = 0$, using the restrictions

F(A) tests $H : \alpha_i$'s all equal

F(B) tests $H : \beta_j$'s all equal.

k. Orthogonal contrasts: Sums of squares SSA and SSB can be partitioned into sums of squares for orthogonal 1-degree-of-freedom contrasts.

Examples

Partitioning SSA

$\alpha_1 - \alpha_2$ $\qquad 3(9 - 11)^2/(1 + 1) \qquad = 6 = \text{SSA}$

Partitioning SSB

$\beta_1 - \beta_3$ $\qquad 2(10 - 9)^2/(1 + 1) \qquad = 1$

$\beta_1 - 2\beta_2 + \beta_3$ $\qquad 2(10 - 22 + 9)^2/(1 + 1 + 2) = \underline{3}$

$4 = \text{SSB}$

Table 12

Analysis of Variance for a 2-way Crossed Classification, 1 Observation Per Cell.

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Square	F-Statistics
General case				
Mean	1	$SSM = ab\bar{y}_{..}^2$	$MSM = SSM$	$F(M) = \frac{MSM}{MSE}$
Plants	a-1	$SSA = b\sum_{i=1}^{a}(\bar{y}_{i.} - \bar{y}_{..})^2$	$MSA = \frac{SSA}{a-1}$	$F(A) = \frac{MSA}{MSE}$
Brands	b-1	$SSB = a\sum_{j=1}^{b}(\bar{y}_{.j} - \bar{y}_{..})^2$	$MSB = \frac{SSB}{b-1}$	$F(B) = \frac{MSB}{MSE}$
Residual	(a-1)(b-1)	$SSE = \sum_{i=1}^{a}\sum_{j=1}^{b}(y_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..})^2$	$MSE = \frac{SSE}{(a-1)(b-1)}$	
Total	ab	$SST = \sum_{i=1}^{a}\sum_{j=1}^{b} y_{ij}^2$		
Example (Table 11)				
Mean	1	SSM = 600	MSM = 600	F(M) = 300
Plants	1	SSA = 6	MSA = 6	F(A) = 1
Brands	2	SSB = 4	MSB = 2	F(B) = 1/3
Residual	2	SSE = 12	MSE = 6	
Total	6	SST = 622		

ℓ. Estimated differences between levels:

$$\hat{\alpha}_i - \hat{\alpha}_{i'} = \bar{y}_{i.} - \bar{y}_{i'.}, \quad \text{for } i \neq i'$$

$$\hat{\beta}_j - \hat{\beta}_{j'} = \bar{y}_{.j} - \bar{y}_{.j'}, \quad \text{for } j \neq j' .$$

m. Variance of an estimated difference between levels:

$$v(\hat{\alpha}_i - \hat{\alpha}_{i'}) = 2\sigma^2/b$$

$$v(\hat{\beta}_j - \hat{\beta}_{j'}) = 2\sigma^2/a .$$

2.4. The 2-way crossed classification with n observations per cell.

a. Example: We use the example of Table 2, concerning 3 brands of loom tested in 5 textile plants, with 2 observations on each brand in each plant. As discussed previously, let

$$y_{ijk} = \text{k'th observation in plant i on brand j} .$$

In general we may have a plants, b brands and n observations on each plant × brand combination. In our case a = 5, b = 3 and n = 2, and we analyze the following data.

Table 13

Data for Table 2

Textile Plant	Brand of Loom A	B	C	Total	Mean
1	7,9	4,8	2,6	36	6
2	9,3	8,6	6,10	42	7
3	1,3	3,3	2,6	18	3
4	2,4	3,7	3,5	24	4
5	8,14	1,7	0,0	30	5
Total	60	50	40	150	
Mean	6	5	4		5

b. Model: The customary linear model is

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk}$$

where μ is a general mean, α_i is the effect due to the i'th plant, β_j is the effect due to the j'th brand, γ_{ij} is the interaction effect [often symbolized as $(\alpha\beta)_{ij}$] due to interaction of the i'th plant and j'th brand, and e_{ijk} is the residual random error assumed to have variance σ^2.

c. Model equations: The model equations for the 6 observations in the first plant are

$$
\begin{aligned}
7 &= \mu + \alpha_1 + \beta_1 + \gamma_{11} + e_{111} \\
9 &= \mu + \alpha_1 + \beta_1 + \gamma_{11} + e_{112} \\
4 &= \mu + \alpha_1 + \beta_2 + \gamma_{12} + e_{121} \\
8 &= \mu + \alpha_1 + \beta_2 + \gamma_{12} + e_{122} \\
2 &= \mu + \alpha_1 + \beta_3 + \gamma_{13} + e_{131} \\
6 &= \mu + \alpha_1 + \beta_3 + \gamma_{13} + e_{132}
\end{aligned}
$$

Those for the other 4 plants follow similarly.

d. Normal equations: The normal equations $\underset{\sim}{X}'\underset{\sim}{X}\hat{\underset{\sim}{b}} = \underset{\sim}{X}'\underset{\sim}{y}$ have the following general form:

$$abn\hat{\mu} + bn\sum_{i=1}^{a}\hat{\alpha}_i + an\sum_{j=1}^{b}\hat{\beta}_j + n\sum_{i=1}^{a}\sum_{j=1}^{b}\hat{\gamma}_{ij} = y_{...}$$

$$bn\hat{\mu} + bn\hat{\alpha}_i + n\sum_{j=1}^{b}\hat{\beta}_j + n\sum_{j=1}^{b}\hat{\gamma}_{ij} = y_{i..}\,,$$

one equation for each $i = 1, \cdots, a$

$$an\hat{\mu} + n\sum_{i=1}^{a}\hat{\alpha}_i + an\hat{\beta}_j + n\sum_{j=1}^{a}\hat{\gamma}_{ij} = y_{.j.}\,,$$

one equation for each $j = 1, \cdots, b$

$$n\hat{\mu} + \hat{\alpha}_i + n\hat{\beta}_j + n\hat{\gamma}_{ij} = y_{ij.}\,,$$

one equation for each combination of $i = 1, \cdots, a$ and $j = 1, \cdots, b$.

It is left as an exercise for the reader to write out these equations in full for the data of the example.

e. Restrictions on model: With balanced data, as here, we usually have the following restrictions:

$$\sum_{i=1}^{a} \alpha_i = 0 \qquad \sum_{i=1}^{a} \gamma_{ij} = 0 \quad \text{for all } j = 1, \cdots, b$$

$$\sum_{j=1}^{b} \beta_j = 0 \qquad \sum_{j=1}^{b} \gamma_{ij} = 0 \quad \text{for all } i = 1, \cdots, a.$$

f. Constraints on solution: In order to get one of the infinitely many solutions to the normal equations we utilize constraints on the solution analogous to the restrictions on the model

$$\sum_{i=1}^{a} \hat{\alpha}_i = 0 \qquad \sum_{i=1}^{a} \hat{\gamma}_{ij} = 0 \quad \text{for all } j = 1, \cdots, b$$

$$\sum_{j=1}^{b} \hat{\beta}_j = 0 \qquad \sum_{j=1}^{b} \hat{\gamma}_{ij} = 0 \quad \text{for all } i = 1, \cdots, a.$$

g. Solution: The solution vector $\hat{\underset{\sim}{b}}$ has elements

$$\hat{\mu} = \bar{y}_{...} \qquad \hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...} \qquad \hat{\gamma}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} - \bar{y}_{.j.} + \bar{y}_{...}$$

$$\hat{\beta}_j = \bar{y}_{.j.} - \bar{y}_{...} \; .$$

We leave it to the reader to derive these values from Table 13.

h. Analysis of variance: The general form of the analysis of variance table is shown in Table 6. With the inclusion of $SSM = abn\bar{y}^2_{...}$ for a first line (as a source of variation due to the mean), in which case the total sum of squares becomes

$$SST = SST_m + SSM = \sum_i \sum_j \sum_k (y_{ijk} - \bar{y}_{...})^2 + abn\bar{y}^2_{...} = \sum_i \sum_j \sum_k y^2_{ijk} \; ,$$

the analysis of variance for the example is as shown in Table 14.

Table 14

Analysis of Variance for Data of Table 13.

Source of Variation	Degrees of Freedom	Sums of Squares	Mean Square	F-Statistics
Mean	1	SSM = 750	MSM = 750	F(M) = 750/7
Plants	4	SSA = 60	MSA = 15	F(A) = 15/7
Brands	2	SSB = 20	MSB = 10	F(B) = 10/7
Interaction	8	SSAB = 132	MSAB = $16\frac{1}{2}$	F(AB) = $16\frac{1}{2}/7$
Residual	15	SSE = 104	MSE = 7 (approx.)	
Total	30	SST = 1066		

i. Estimated residual variance:

$$\hat{\sigma}^2 = \text{MSE} = \frac{\sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}(y_{ijk} - \bar{y}_{ij.})^2}{ab(n-1)} = 6\frac{14}{15} = 7 \text{ (approx.)}.$$

j. Hypothesis testing using F's: In the presence of the restrictions included in the model, the F-statistics of the analysis of variance table can be used for hypothesis testing as follows:

F(M)	tests	$H: \mu = 0$, i.e., $H: E(\bar{y}_{...}) = 0$, using the restrictions
F(A)	tests	$H:$ all α_i equal
F(B)	tests	$H:$ all β_j equal
F(AB)	tests	$H:$ all γ_{ij} equal.

k. Orthogonal contrasts: Sums of squares SSA, SSB and SSAB can be partitioned into sums of squares for orthogonal 1-degree-of-freedom contrasts.

Examples

Partitioning SSA

$\alpha_1 - \alpha_2$	$6(6-7)^2/(1+1)$	$= 3$
$\alpha_1 + \alpha_2 - 2\alpha_3$	$6(6+7-6)^2/(1+1+4)$	$= 49$
$\alpha_1 + \alpha_2 + \alpha_3 - 3\alpha_4$	$6(6+7+3-12)^2/(1+1+1+9)$	$= 8$
$\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 - 4\alpha_5$	$6(6+7+3+4-20)^2/(1+1+1+1+16)$	$= \underline{0}$
		$\underline{60} = SSA$

Partitioning SSB

$\beta_1 - \beta_2$	$10(6-5)^2/(1+1)$	$= 5$
$\beta_1 + \beta_2 - 2\beta_3$	$10(6+5-8)^2/(1+1+4)$	$= \underline{15}$
		$\underline{20} = SSB$

ℓ. Estimated differences between levels:

$$\hat{\alpha}_i - \hat{\alpha}_{i'} = \bar{y}_{i..} - \bar{y}_{i'..} \quad \text{for } i \neq i'$$

$$\hat{\beta}_j - \hat{\beta}_{j'} = \bar{y}_{.j.} - \bar{y}_{.j'.} \quad \text{for } j \neq j' \; .$$

m. Variance of an estimated difference between levels:

$$v(\hat{\alpha}_i - \hat{\alpha}_{i'}) = 2\sigma^2/bn$$

$$v(\hat{\beta}_j - \hat{\beta}_{j'}) = 2\sigma^2/an \; .$$

Comment

It is clear that when n = 1 the results of this section reduce to those of the preceding one.

CHAPTER 3

3. GENERAL LINEAR MODEL THEORY: A BRIEF SUMMARY

This is the only chapter of these notes which is mostly theory. It is a slight expansion of the summary in LM 227-229.

3.1. Numerical example

A simple numerical example is used, without comment, to illustrate features of the general theory. It consists of 6 observations made in 3 plants, three in the first, two in the second and one in the third.

	Plant 1	Plant 2	Plant 3		
	$y_{11} = 101$	$y_{21} = 84$	$y_{31} = 32$		[LM 165]
	$y_{12} = 105$	$y_{22} = 88$			
	$y_{13} = 94$				
Totals:	$y_{1.} = 300$	$y_{2.} = 172$	$y_{3.} = 32$	$y_{..} = 504$	

3.2. Normal equations and their solution

a. Model

Theory	Example	
$\underset{\sim}{y}'$ = data as a row vector	$\underset{\sim}{y}' = [101 \quad 105 \quad 94 \quad 84 \quad 88 \quad 32]$	[LM 166]
$E(\underset{\sim}{y}) = \underset{\sim}{Xb}$	$E(y_{ij}) = \mu + \alpha_i$	
$\underset{\sim}{b}'$ = row vector of parameters	$\underset{\sim}{b}' = [\mu \quad \alpha_1 \quad \alpha_2 \quad \alpha_3]$	
$\underset{\sim}{X}$ = known matrix, often of 0's and 1's, then called incidence or design matrix		[LM 166]

Theory	Example

$\underset{\sim}{e}$ = vector of residuals

$$\overset{\text{def}}{\equiv} \underset{\sim}{y} - E(\underset{\sim}{y})$$

$$= \underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b}$$

$$\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$$

$$\underset{\sim}{y} = \begin{bmatrix} 101 \\ 105 \\ 94 \\ 84 \\ 88 \\ 32 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \underset{\sim}{e}$$ [LM 164]

$$E(\underset{\sim}{e}) = E(\underset{\sim}{y}) - E(\underset{\sim}{y}) = \underset{\sim}{0}$$

$$\text{var}(\underset{\sim}{e}) \overset{\text{def}}{\equiv} \sigma^2 \underset{\sim}{I}$$

$$\underset{\sim}{e} \sim (\underset{\sim}{0}, \sigma^2 \underset{\sim}{I})$$

$$\underset{\sim}{y} \sim (\underset{\sim}{X}\underset{\sim}{b}, \sigma^2 \underset{\sim}{I})$$ [LM 166]

b. Least squares

Minimize $\sum_t [y_t - E(y_t)]^2$

$$= [\underset{\sim}{y} - E(\underset{\sim}{y})]'[\underset{\sim}{y} - E(\underset{\sim}{y})]$$

$$= \underset{\sim}{e}'\underset{\sim}{e}$$

$$= (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b})'(\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{b})$$

This yields normal equations. [LM 165, 80]

c. Normal equations

$$\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^{o} = \underset{\sim}{X}'\underset{\sim}{y}$$

$$\begin{bmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu^o \\ \alpha_1^o \\ \alpha_2^o \\ \alpha_3^o \end{bmatrix} = \begin{bmatrix} 504 \\ 300 \\ 172 \\ 32 \end{bmatrix}$$ [(6), LM 168]

Theory | Example

d. Generalized inverses of $\underset{\sim}{X}'\underset{\sim}{X}$

$$\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{X}'\underset{\sim}{X} \qquad [(1),\ \text{LM } 1]$$

Many $\underset{\sim}{G}$'s [LM 2]

For all of them

$$\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}'\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{X}'\underset{\sim}{X}$$

$$\underset{\sim}{X}\underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{X} \qquad [\text{Th. } 7,\ \text{LM } 20]$$

$\underset{\sim}{X}\underset{\sim}{G}\underset{\sim}{X}'$ symmetric

$\underset{\sim}{X}\underset{\sim}{G}\underset{\sim}{X}'$ invariant to $\underset{\sim}{G}$

Define $\underset{\sim}{H} = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X}$, idempotent. [LM 169]

e. A solution of normal equations

$$\underset{\sim}{b}^{o} = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{y} \qquad [(7),\ \text{LM } 168]$$

$$E(\underset{\sim}{b}^{o}) = \underset{\sim}{H}\underset{\sim}{b} \qquad [(8),\ \text{LM } 169]$$

$$\text{var}(\underset{\sim}{b}^{o}) = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}'\sigma^2 \qquad [(9),\ \text{LM } 169]$$

$$\underset{\sim}{b}^{o} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 504 \\ 300 \\ 172 \\ 32 \end{bmatrix} = \begin{bmatrix} 0 \\ 100 \\ 86 \\ 32 \end{bmatrix} \qquad [(22),\ \text{LM } 172]$$

3.3. Sums of squares

a. Predicted $\underset{\sim}{y}$, i.e., estimating $E(\underset{\sim}{y})$

$$\hat{\underset{\sim}{y}} = \widehat{E(\underset{\sim}{y})} = \underset{\sim}{X}\underset{\sim}{b}^{o} = \underset{\sim}{X}\underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{y} \qquad [(10),\ \text{LM } 170]$$

Theory	Example
Note: $\underset{\sim}{\hat{y}}$ is invariant to $\underset{\sim}{G}$	$\begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 100 \\ 86 \\ 32 \end{bmatrix} = \begin{bmatrix} 100 \\ 100 \\ 100 \\ 86 \\ 86 \\ 32 \end{bmatrix}$ [(23), LM 172]

b. Residual sum of squares

$$SSE = \sum_t (y_t - \hat{y}_t)^2$$

$$= (\underset{\sim}{y} - \underset{\sim}{\hat{y}})'(\underset{\sim}{y} - \underset{\sim}{\hat{y}})$$

$$= \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{b}^{o\prime}\underset{\sim}{X}'\underset{\sim}{y} \qquad [(12),\ LM\ 170]$$

$$= \underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{XGX'})\underset{\sim}{y} \qquad [(11),\ LM\ 170]$$

Note: (12) is ideal for computation.

(11) is useful for theory.

SSE is invariant to $\underset{\sim}{G}$: see (11).

$SST = \Sigma y_i^2 = \underset{\sim}{y}'\underset{\sim}{y}$ $\qquad$ $SST = 101^2 + 105^2 + 94^2 + 84^2 + 88^2 + 32^2 = 45,886$

$SSR = SST - SSE$

$= \underset{\sim}{y}'\underset{\sim}{XGX'}\underset{\sim}{y}$

$= \underset{\sim}{b}^{o\prime}\underset{\sim}{X}'\underset{\sim}{y}$ $\qquad$ $SSR = 0(504) + 100(300) + 86(172) + 32(32) = 45,816$

$SSE = SST - SSR$ $\qquad = 70$

c. Partitioning of sums of squares [LM 172-3]

$$SSR \begin{cases} SSM = N\bar{y}^2 = \underset{\sim}{y}'\underset{\sim}{N}^{-1}\underset{\sim}{\mathbf{11}}'\underset{\sim}{y} & = 42,336 \\ SSR_m = \underset{\sim}{y}'(\underset{\sim}{XGX'} - \underset{\sim}{N}^{-1}\underset{\sim}{\mathbf{11}}')\underset{\sim}{y} & = 3,480 \end{cases}$$

$$SSE = \underset{\sim}{y}'(\underset{\sim}{I} - \underset{\sim}{XGX'})\underset{\sim}{y} = 70$$

$$SST = \underset{\sim}{y}'\underset{\sim}{y} = 45,886$$

[Tables 5.3 and 5.4 LM 171, 173]

3.4. Analysis of variance

a. Table 5.6a LM 177

Analysis of Variance for Fitting the Model

$$\underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{e}$$

Source of Variation[2]	d.f.[1]	Sum of Squares	Mean Square	F-Statistics
Mean	1	$SSM = N\bar{y}^2$	$MSM = \frac{SSM}{1}$	$F(M) = \frac{MSM}{MSE}$
Model (a.f.m.)	r - 1	$SSR_m = \underset{\sim}{b}^{o\prime}\underset{\sim}{X}'\underset{\sim}{y} - N\bar{y}^2$	$MSR_m = \frac{SSR_m}{r-1}$	$F(R_m) = \frac{MSR_m}{MSE}$
Residual error	N - r	$SSE = \underset{\sim}{y}'\underset{\sim}{y} - \underset{\sim}{b}^{o\prime}\underset{\sim}{X}'\underset{\sim}{y}$	$MSE = \frac{SSE}{N-r}$	
Total	N	$SST = \underset{\sim}{y}'\underset{\sim}{y}$		

[1] $r = r(\underset{\sim}{X})$.

[2] a.f.m. = after fitting the mean.

b. Table 5.7 LM 179

Table 5.6a for the Example

Source of Variation	d.f.	Sum of Squares	Mean Square	F-Statistics
Mean	1	SSM = 42,336	42,336	F(M) = 1814.4
Model (a.f.m.)	2	SSR_m = 3,480	1,740	$F(R_m)$ = 74.3
Residual error	3	SSE = 70	$23\frac{1}{3}$	
Total	6	SST = 45,886		

c. Estimating σ^2

$$\hat{\sigma}^2 = \frac{SSE}{N - r(\underset{\sim}{X})} = MSE = \frac{70}{3}.$$

Note: No distributional assumptions needed up to this point: only $\underset{\sim}{e} \sim (\underset{\sim}{0}, \sigma^2\underset{\sim}{I})$.

d. F-statistics [LM 178-180]

Now assume normality: that $\underset{\sim}{y} \sim (\underset{\sim}{Xb},\ \sigma^2\underset{\sim}{I})$ is $\underset{\sim}{y} \sim N(\underset{\sim}{Xb},\ \sigma^2\underset{\sim}{I})$.

$$F(R) = \frac{SSR/r(\underset{\sim}{X})}{MSE} \quad \text{tests} \quad H: \underset{\sim}{Xb} = \underset{\sim}{0}, \quad \text{using } F_{r(\underset{\sim}{X}),N-r(\underset{\sim}{X})}$$

Note: The hypothesis is $\underset{\sim}{Xb} = \underset{\sim}{0}$, not $\underset{\sim}{b} = \underset{\sim}{0}$.

$$F(M) = \frac{SSM}{MSE} = \frac{N\bar{y}^2}{\hat{\sigma}^2} = \left(\frac{\bar{y}}{\hat{\sigma}/\sqrt{N}}\right)^2 = t^2, \qquad \text{using } F_{1,N-r(\underset{\sim}{X})}\,.$$

This tests $H: E(\bar{y}) = 0$, i.e., $\underset{\sim}{1}'\underset{\sim}{Xb} = 0$.

Notes: This is the F for "testing the mean".
It is not always "$H: \mu = 0$".
The F is the square of the usual "t".

$$F(R_m) = \frac{SSR_m/[r(\underset{\sim}{X}) - 1]}{MSE} \quad \text{using } F_{r(\underset{\sim}{X})-1,N-r(\underset{\sim}{X})} \text{ tests concordance of}$$

the data with the model $E(\underset{\sim}{y}) = \underset{\sim}{Xb}$ over and above the mean.

3.5. Estimable functions

a. Example

The normal equations $\underset{\sim}{X}'\underset{\sim}{Xb}^o = \underset{\sim}{X}'\underset{\sim}{y}$, e.g.

$$\begin{bmatrix} 6 & 3 & 2 & 1 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu^o \\ \alpha_1^o \\ \alpha_2^o \\ \alpha_3^o \end{bmatrix} = \begin{bmatrix} 504 \\ 300 \\ 172 \\ 32 \end{bmatrix},$$

have many solutions. We give five; and below each is the corresponding value of certain functions of the elements of the solutions.

Five Solutions [LM 160]

Element	1	2	3	4	5	
μ^o	0	84	$72\frac{3}{4}$	32	5,283	
α_1^o	100	16	$27\frac{1}{4}$	68	-5,183	
α_2^o	86	2	$13\frac{1}{4}$	54	-5,197	
α_3^o	32	-52	$-40\frac{3}{4}$	0	-5,251	
Functions						
$\alpha_1^o+\alpha_2^o$	186	18	$40\frac{1}{2}$	122	-10,380	- all different
$(\alpha_1^o+\alpha_2^o+\alpha_3^o)/3$	73	$-11\frac{1}{3}$	-1/12	$40\frac{2}{3}$	$-5,210\frac{1}{3}$	- all different
$\alpha_1^o-\alpha_2^o$	14	14	14	14	14	- all the same
$\mu+\alpha_1^o$	100	100	100	100	100	- all the same
$\mu+\alpha_3^o$	32	32	32	32	32	- all the same

b. Definition

In all cases of the general linear model $E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b}$, with normal equations $\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^o = \underset{\sim}{X}'\underset{\sim}{y}$, there are certain linear functions of the elements of the solution vector, $\underset{\sim}{q}'\underset{\sim}{b}^o$, that are invariant to whatever solution $\underset{\sim}{b}^o$ is used — functions like the last three in the preceding example. They each have the property that $\underset{\sim}{q}' = \underset{\sim}{t}'\underset{\sim}{X}$ for some $\underset{\sim}{t}'$. Corresponding to such a function $\underset{\sim}{q}'\underset{\sim}{b}^o$ is the similar function $\underset{\sim}{q}'\underset{\sim}{b}$ of the parameter vector: it is called an estimable function. It has the following properties:

$$\underset{\sim}{q}' = \underset{\sim}{t}'\underset{\sim}{X} \qquad [(40),\ \text{LM } 181]$$

$\underset{\sim}{q}'\underset{\sim}{b}$ is <u>estimable</u>

$\underset{\sim}{q}'\underset{\sim}{b}^o$ is invariant to $\underset{\sim}{b}^o$ [LM 181]

$$E(\underset{\sim}{q}'\underset{\sim}{b}^o) = \underset{\sim}{q}'\underset{\sim}{b}$$

$$v(\underset{\sim}{q}'\underset{\sim}{b}^o) = \underset{\sim}{q}'\underset{\sim}{G}\underset{\sim}{q}\sigma^2 \qquad [(43),\ \text{LM } 182]$$

The b.l.u.e. of $\underset{\sim}{q}'\underset{\sim}{b}$ is $\widehat{\underset{\sim}{q}'\underset{\sim}{b}} = \underset{\sim}{q}'\underset{\sim}{b}^o$. [(41), LM 181] b.l.u.e. means <u>b</u>est <u>l</u>inear <u>u</u>nbiased <u>e</u>stimator. [LM 182]

$\underset{\sim}{q}'\underset{\sim}{b}^o$ is <u>linear</u> in $\underset{\sim}{y}$ because $\underset{\sim}{q}'\underset{\sim}{b}^o = \underset{\sim}{q}'\underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{y} = (\underset{\sim}{q}'\underset{\sim}{G}\underset{\sim}{X}')\underset{\sim}{y}$

$\underset{\sim}{q}'\underset{\sim}{b}^o$ is <u>unbiased</u> because $E(\underset{\sim}{q}'\underset{\sim}{b}^o) = \underset{\sim}{q}'\underset{\sim}{b}$

$\underset{\sim}{q}'\underset{\sim}{b}^o$ is <u>best</u> because, among all linear unbiased estimators of $\underset{\sim}{q}'\underset{\sim}{b}$, it has the smallest variance; i.e.,

$$v(\underset{\sim}{q}'\underset{\sim}{b}^o) = \underset{\sim}{q}'\underset{\sim}{G}\underset{\sim}{q}\sigma^2 \le v(\text{any other linear, unbiased estimator of } \underset{\sim}{q}'\underset{\sim}{b}).$$

c. <u>Useful properties</u>

(i) Expected value of an observation is estimable; [LM 181]

e.g., $E(y_i) = \mu + \alpha_i \Rightarrow \mu + \alpha_i$ is estimable.

(ii) Linear combinations of estimable functions are estimable; [LM 181]

e.g., $\mu + \alpha_1 - (\mu + \alpha_2) \equiv \alpha_1 - \alpha_2$ is estimable.

(iii) A set of linearly independent (LIN) estimable functions cannot contain more than $r(\underset{\sim}{X})$ such functions; [LM 185]

e.g., in the example, $r(\underset{\sim}{X}) = 3$, and so one cannot have more than 3 LIN estimable functions. One possible set is $\mu + \alpha_1$, $\alpha_1 - \alpha_2$, $\alpha_1 - \alpha_3$. All other estimable functions in this example are linear combinations of these three.

(iv) The function $\underset{\sim}{q}'\underset{\sim}{b}$ is estimable if and only if $\underset{\sim}{q}'\underset{\sim}{H} = \underset{\sim}{q}'$; [LM 185]

e.g., $\mu + (\alpha_1 + \alpha_2 + \alpha_3)/3 = (1 \;\; \tfrac{1}{3} \;\; \tfrac{1}{3} \;\; \tfrac{1}{3})\underset{\sim}{b}$

is estimable because

$$[1 \;\; \tfrac{1}{3} \;\; \tfrac{1}{3} \;\; \tfrac{1}{3}]\begin{bmatrix}0&0&0&0\\0&\tfrac{1}{3}&0&0\\0&0&\tfrac{1}{2}&0\\0&0&0&1\end{bmatrix}\begin{bmatrix}6&3&2&1\\3&3&0&0\\2&0&2&0\\1&0&0&1\end{bmatrix} = [1 \;\; \tfrac{1}{3} \;\; \tfrac{1}{3} \;\; \tfrac{1}{3}]\begin{bmatrix}0&0&0&0\\1&1&0&0\\1&0&1&0\\1&0&0&1\end{bmatrix} = [1 \;\; \tfrac{1}{3} \;\; \tfrac{1}{3} \;\; \tfrac{1}{3}].$$

3.6. The general linear hypothesis

a. General form of hypothesis [LM 185]

$H : \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$.

Limitations: $\underset{\sim}{K}'\underset{\sim}{b}$ estimable $\Rightarrow$ $\underset{\sim}{K}' = \underset{\sim}{T}'\underset{\sim}{X}$ for some $\underset{\sim}{T}'$

$\underset{\sim}{K}'$ full row rank $r(\underset{\sim}{K}'_{s \times p}) = s$

$\underset{\sim}{m}$ represents pre-assigned values

$\underset{\sim}{m}$ is often, but not always, $\underset{\sim}{0}$.

Note: $\text{var}(\underset{\sim}{K}'\underset{\sim}{b}^{o}) = \underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K}\sigma^2$.

b. F-statistic

Based on likelihood ratio statistic.

$$F(H) = \frac{(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})}{s\hat{\sigma}^2} = \frac{Q}{s\hat{\sigma}^2}. \qquad [(70),\ \text{LM } 190]$$

$Q = (\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})$.

$(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}$ always exists for $\underset{\sim}{K}' = \underset{\sim}{T}'\underset{\sim}{X}$ of full row rank.

Under $H : \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$, $F(H) \sim F_{s, N-r(\underset{\sim}{X})}$.

Example

$H : \alpha_1 - \alpha_2 = 10.$ [LM 196]

$$\underset{\sim}{K}'\underset{\sim}{b} = \alpha_1 - \alpha_2$$

$$[0 \;\; 1 \;\; -1 \;\; 0]\begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \alpha_1 - \alpha_2$$

$$\underset{\sim}{K}' = [0 \;\; 1 \;\; -1 \;\; 0] \qquad \underset{\sim}{m} = 10$$

$$\mathbf{K}'\mathbf{b}^{\circ} - \mathbf{m} = [0 \quad 1 \ -1 \quad 0]\begin{bmatrix} 0 \\ 100 \\ 86 \\ 32 \end{bmatrix} - 10 = 14 - 10 = 4$$

$$\mathbf{K}'\mathbf{G}\mathbf{K} = [0 \quad 1 \ -1 \quad 0]\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \end{bmatrix} = \frac{1}{3} + \frac{1}{2} = \frac{5}{6}$$

$$F = \frac{4(5/6)^{-1}4}{1(70/3)} = \frac{3(6)16}{5(70)} = \frac{144}{175} .$$

c. Maximum number of statements in hypothesis

For testable hypotheses $\mathbf{K}' = \mathbf{T}'\mathbf{X}$, so that $r(\mathbf{K}') = s \leq r(\mathbf{X}) = r$. Therefore in $H: \mathbf{K}'\mathbf{b} = \mathbf{m}$ there cannot be more than $r(\mathbf{X})$ functions in $\mathbf{K}'\mathbf{b}$; i.e., there cannot be more than $r(\mathbf{X})$ rows in $\mathbf{K}'$.

Furthermore, when in the hypothesis $H: \mathbf{K}'\mathbf{b} = \mathbf{0}$ (i.e., $\mathbf{m} \equiv \mathbf{0}$), we have $r(\mathbf{K}') = r(\mathbf{X})$ then $Q = SSR$.

Proof (adapted from W. H. Swallow):

$$\begin{aligned} Q &= (\mathbf{K}'\mathbf{b}^{\circ} - \mathbf{m})'(\mathbf{K}'\mathbf{G}\mathbf{K})^{-1}(\mathbf{K}'\mathbf{b}^{\circ} - \mathbf{m}) \\ &= (\mathbf{K}'\mathbf{b}^{\circ})'(\mathbf{K}'\mathbf{G}\mathbf{K})^{-1}\mathbf{K}'\mathbf{b}^{\circ} \quad \text{for } \mathbf{m} = \mathbf{0} \\ &= \mathbf{y}'\mathbf{X}\mathbf{G}'\mathbf{K}(\mathbf{K}'\mathbf{G}\mathbf{K})^{-1}\mathbf{K}'\mathbf{G}\mathbf{X}'\mathbf{y} . \end{aligned}$$

Because $\mathbf{X}'\mathbf{X}$ is symmetric of rank r, there exists a full column rank matrix $\mathbf{M}$ such that $\mathbf{X}'\mathbf{X} = \mathbf{M}\mathbf{M}'$ and $(\mathbf{M}'\mathbf{M})^{-1}$ exists. Also, the Penrose inverse of $\mathbf{X}'\mathbf{X}$ is $\mathbf{M}(\mathbf{M}'\mathbf{M})^{-2}\mathbf{M}$, and using it for $\mathbf{G}$ in Q gives

$$\begin{aligned} Q &= \mathbf{y}'\mathbf{X}\mathbf{M}(\mathbf{M}'\mathbf{M})^{-2}\mathbf{M}'\mathbf{K}[\mathbf{K}'\mathbf{M}(\mathbf{M}'\mathbf{M})^{-2}\mathbf{M}'\mathbf{K}]^{-1}\mathbf{K}'\mathbf{M}(\mathbf{M}'\mathbf{M})^{-2}\mathbf{M}'\mathbf{X}'\mathbf{y} \\ &= \mathbf{y}'\mathbf{X}\mathbf{M}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{L}'(\mathbf{L}\mathbf{L}')^{-1}\mathbf{L}(\mathbf{M}'\mathbf{M})^{-1}\mathbf{M}\mathbf{X}'\mathbf{y} \quad \text{for } \mathbf{L} \equiv \mathbf{K}'\mathbf{M}(\mathbf{M}'\mathbf{M})^{-1} . \end{aligned}$$

Because $\underset{\sim}{K}'$ has full row rank r and $\underset{\sim}{M}$ has full column rank r, $\underset{\sim}{L}$ is square, and we see that $(\underset{\sim}{LL}')^{-1}$ exists. Therefore $|\underset{\sim}{L}| \neq \underset{\sim}{0}$ and so $\underset{\sim}{L}^{-1}$ exists. Hence

$$Q = \underset{\sim}{y}'\underset{\sim}{XM}(\underset{\sim}{M}'\underset{\sim}{M})^{-2}\underset{\sim}{MX}'\underset{\sim}{y} = \underset{\sim}{y}'\underset{\sim}{XGX}'\underset{\sim}{y} = SSR\,.$$

Q.E.D.

d. Estimation under the hypothesis

Under the hypothesis $H : \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$ we have:

$$\underset{\sim}{b}^{o}_{H} = \underset{\sim}{b}^{o} - \underset{\sim}{GK}(\underset{\sim}{K}'\underset{\sim}{GK})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m}) \qquad [(72),\ \text{LM } 191]$$

$$(\underset{\sim}{y} - \underset{\sim}{Xb}^{o}_{H})'(\underset{\sim}{y} - \underset{\sim}{Xb}^{o}_{H}) = SSE_H = SSE + Q \qquad [(74),\ \text{LM } 191]$$

$$\hat{\sigma}^2_H = SSE_H/[N - r(\underset{\sim}{X}) - s].$$

e. Calculation of numerator sum of squares

The numerator sum of squares is $Q = (\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{GK})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{o} - \underset{\sim}{m})$.

$$Q = SSE_H - SSE$$

Full model: $\underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{e}$ $\qquad SSE \equiv SSE_{full}$

Reduced model: $\underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{e}$ and $\underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$ $\qquad SSE_H = SSE_{red}$

$$Q = SSE_{red} - SSE_{full}\,.$$

Although $SSE = \underset{\sim}{y}'\underset{\sim}{y} - SSR$, it is to be observed that

$$Q \neq \underset{\sim}{y}'\underset{\sim}{y} - SSR_{red} - (\underset{\sim}{y}'\underset{\sim}{y} - SSR_{full})$$

$$\neq SSR_{full} - SSR_{red}$$

because $\underset{\sim}{y}'\underset{\sim}{y}$ is not necessarily the total sum of squares for the reduced model and so $SSE_{red} \neq \underset{\sim}{y}'\underset{\sim}{y} - SSR_{red}$.

Example [LM 117]

Full model: $y_i = b_0 + b_1x_{i1} + b_2x_{i2} + e_i$

Reduced model: the above and $b_1 = b_2 + 4$;

$$\text{i.e.,} \quad y_i = b_0 + (b_2 + 4)x_{i1} + b_2x_{i2} + e_i$$

$$y_i - 4x_{i1} = b_0 + b_2(x_{i1} + x_{i2}) + e_i$$

$$SST_{red} = (\underset{\sim}{y} - 4\underset{\sim}{x}_1)'(\underset{\sim}{y} - 4\underset{\sim}{x}_1) \neq \underset{\sim}{y}'\underset{\sim}{y} .$$

f. Special case

$$H : \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{0} \; ; \text{ i.e., } \underset{\sim}{m} \equiv \underset{\sim}{0} .$$

Always: $Q = SSE_{red} - SSE_{full}$

For $\underset{\sim}{m} \equiv \underset{\sim}{0}$: $Q = SSR_{full} - SSR_{red}$.

Analyses of variance and F-tests: See LM 192-193.

g. Testing non-testable hypotheses - (LM 193-199, corrected in paper BU-501-M, which is in the appendix to these notes.)

In $F = Q/s\hat{\sigma}^2$ with $Q = (\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})$, we always use a $\underset{\sim}{K}'$ of full row rank. This and the estimability of $\underset{\sim}{K}'\underset{\sim}{b}$ ensure that F can be calculated; i.e., that $(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}$ exists. Thus estimability of $\underset{\sim}{K}'\underset{\sim}{b}$ is a sufficient condition for the existence of $(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}$; but it is not a necessary condition. Hence there are values of $\underset{\sim}{K}'$ for which $(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}$ exists without $\underset{\sim}{K}'\underset{\sim}{b}$ being estimable. The question then is "What hypothesis *is* F testing?"

We have two situations. In both of them $\underset{\sim}{G}$ must be symmetric and reflexive: $\underset{\sim}{G} = \underset{\sim}{G}' = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}$. If it is not, use $\underset{\sim}{G}^* = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}$, which is. Then $\underset{\sim}{H}^* = \underset{\sim}{G}^*\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{H}$ and $\underset{\sim}{b}^* = \underset{\sim}{G}^*\underset{\sim}{X}'\underset{\sim}{y} = \underset{\sim}{b}^o$, and calculate F using $(\underset{\sim}{K}'\underset{\sim}{G}^*\underset{\sim}{K})^{-1}$.

$\underset{\sim}{K}'\underset{\sim}{b}$ not estimable, $(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}$ existing, F calculable

$$F \quad \text{tests} \quad H: \underset{\sim}{K}'\underset{\sim}{H}\underset{\sim}{b} = \underset{\sim}{m} .$$

$\underset{\sim}{K}'\underset{\sim}{b}$ partially estimable, $(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}$ existing, F calculable

Partition

$$\underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m} \quad \text{as} \quad \begin{bmatrix} \underset{\sim}{K}_1'\underset{\sim}{b} \\ \underset{\sim}{K}_2'\underset{\sim}{b} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{m}_1 \\ \underset{\sim}{m}_2 \end{bmatrix}, \quad \begin{matrix} \underset{\sim}{K}_1'\underset{\sim}{b} \text{ estimable} \\ \underset{\sim}{K}_2'\underset{\sim}{b} \text{ non-estimable} \end{matrix} .$$

Then

$$F \quad \text{tests} \quad H: \begin{bmatrix} \underset{\sim}{K}_1'\underset{\sim}{b} \\ \underset{\sim}{K}_2'\underset{\sim}{H}\underset{\sim}{b} \end{bmatrix} = \begin{bmatrix} \underset{\sim}{m}_1 \\ \underset{\sim}{m}_2 \end{bmatrix} .$$

h. Independent and orthogonal contrasts [LM 199-204]

For $H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{0}$ with $r(\underset{\sim}{K}') = s$, the F-statistic for testing

$$H: \text{simultaneously} \left\{ \begin{matrix} H_1 : \underset{\sim}{k}_1'\underset{\sim}{b} = 0 \\ H_2 : \underset{\sim}{k}_2'\underset{\sim}{b} = 0 \\ \vdots \\ H_s : \underset{\sim}{k}_s'\underset{\sim}{b} = 0 \end{matrix} \right\} \text{is } F(H) = Q/s\hat{\sigma}^2 .$$

Suppose we test each H_i individually

$$H_i : \underset{\sim}{k}_i'\underset{\sim}{b} = 0 \quad \text{with} \quad F(H_i) = q_i/\hat{\sigma}^2$$

$$q_i = \underset{\sim}{b}^{o\prime}\underset{\sim}{k}_i(\underset{\sim}{k}_i'\underset{\sim}{G}\underset{\sim}{k}_i)^{-1}\underset{\sim}{k}_i'\underset{\sim}{b}^o = \frac{(\underset{\sim}{k}'\underset{\sim}{b}^o)^2}{\underset{\sim}{k}_i'\underset{\sim}{G}\underset{\sim}{k}_i} .$$

Theorem: The q_i are independent if and only if $\underset{\sim}{k}_i'\underset{\sim}{G}\underset{\sim}{k}_j = 0$ for all $i \neq j = 1, 2, \cdots, s$; and then

$$Q = \sum_{i=1}^{s} q_i .$$

Definition: We say $\underset{\sim}{k}_i'\underset{\sim}{b}$ and $\underset{\sim}{k}_j'\underset{\sim}{b}$ are orthogonal when $\underset{\sim}{k}_i'\underset{\sim}{G}\underset{\sim}{k}_j = 0$.

Note 1: With balanced data, $\underset{\sim}{G}$ is often a scalar matrix, or else it and the $\underset{\sim}{k}_i$'s of interest partition in such a way that the condition reduces to $\underset{\sim}{k}_i^{*\prime}\underset{\sim}{k}_j^{*} = 0$ where $\underset{\sim}{k}_i^{*}$ is a sub-vector of $\underset{\sim}{k}_i$.

Note 2 (with acknowledgment to W. T. Federer): This is a theorem giving the necessary and sufficient condition for independence of q_i and q_j. If they are all pairwise independent then they "add up", i.e., $\Sigma q_i = Q$. It is not a theorem giving a necessary and sufficient condition for "adding up". Independence is sufficient for this, but not necessary; i.e., it is possible to have q_i's for which $\Sigma q_i = Q$, but where the q_i's are not independent. A necessary condition for "adding up" is not known.

3.7. Restricted models

Definitions

Unrestricted model: $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$.

Restricted model: $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$

$$\underset{\sim}{P}'\underset{\sim}{b} = \underset{\sim}{\delta} .$$

Details are given in LM 204-209.

There is a summary in paper BU-533-M, in the appendix to these notes.

Two cases of the restricted model:

(i) $\underset{\sim}{P}'\underset{\sim}{b}$ estimable: section (a) in LM 206

section (b) in BU-533-M

Correction to LM 207 is given in BU-451-M.

(ii) $\underset{\sim}{P}'\underset{\sim}{b}$ non-estimable: section (b) in LM 208

section (a) in BU-533-M.

Papers BU-451-M and BU-533-M are included in the appendix to these notes.

3.8. Constraints on solutions

Details in LM 209-220.

Summary on LM 215 and in BU-533-M (top page 5).

3.9. Generalized least squares

$$\underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{e} \sim (\underset{\sim}{Xb}, \underset{\sim}{V})$$

a. Non-singular $\underset{\sim}{V}$ [LM 220-221]

GLS (or Aitken) equations.

$$\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{Xb}^{o} = \underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$$

$$\underset{\sim}{b}^{o} = (\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$$

$$SSE = \underset{\sim}{y}'\underset{\sim}{V}^{-1}\underset{\sim}{y} - \underset{\sim}{b}^{o}{}'\underset{\sim}{X}'\underset{\sim}{V}^{-1}\underset{\sim}{y}$$

$\underset{\sim}{q}'\underset{\sim}{b}$ is estimable if $\underset{\sim}{q}' = \underset{\sim}{t}'\underset{\sim}{X}$ for some $\underset{\sim}{t}'$.

b. Singular $\underset{\sim}{V}$ [LM 221-223]

$$\underset{\sim}{X}'\underset{\sim}{V}^{-}\underset{\sim}{Xb}^{o} = \underset{\sim}{X}'\underset{\sim}{V}^{-}\underset{\sim}{y}$$

$$\underset{\sim}{b}^{o} = (\underset{\sim}{X}'\underset{\sim}{V}^{-}\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{V}^{-}\underset{\sim}{y}$$

$$SSE = \underset{\sim}{y}'\underset{\sim}{V}^{-}\underset{\sim}{y} - \underset{\sim}{b}^{o}{}'\underset{\sim}{X}'\underset{\sim}{V}^{-}\underset{\sim}{y}$$

$\underset{\sim}{q}'\underset{\sim}{b}$ is estimable if $\underset{\sim}{q}' = \underset{\sim}{t}'\underset{\sim}{M}'\underset{\sim}{X}$ for $\underset{\sim}{V} = \underset{\sim}{MM}'$.

3.10. The R( )-notation

[LM 246-249]

a. Definition

The reduction in sum of squares SSR for fitting the model

$$y = Xb + e\,.$$

is usefully given the symbol $R(b)$:

$$R(b) \equiv \text{SSR for } E(y) = Xb$$

$$R(b) = y'XGX'y = y'X(X'X)^{-}X'y = b^{o'}X'y$$

b. Partitioned models

$$y = X_1b_1 + X_2b_2 + e$$

$$R(b_1, b_2) = y'[X_1 \quad X_2]\begin{bmatrix} X_1'X_1 & X_1'X_2 \\ X_2'X_1 & X_2'X_2 \end{bmatrix}^{-}\begin{bmatrix} X_1' \\ X_2' \end{bmatrix}y$$

$$R(b_1) = y'X_1(X_1'X_1)^{-}X_1'y$$

$$R(b_2) = y'X_2(X_2'X_2)^{-}X_2'y$$

Definitions: $R(b_1|b_2) = R(b_1, b_2) - R(b_2)$

$$R(b_2|b_1) = R(b_1, b_2) - R(b_1)$$

It can be shown that

$$R(b_1|b_2) = y'T_1X_2(X_2'T_1X_2)^{-}X_2'T_1y \quad \text{for} \quad T_1 = I - X_1(X_1'X_1)^{-}X_1'$$

$$R(b_2|b_1) = y'T_2X_1(X_1'T_2X_1)^{-}X_1'T_2y \quad \text{for} \quad T_2 = I - X_2(X_2'X_2)^{-}X_2'\,.$$

c. Examples in specific models

Model	SSR	
$E(y_i) = \mu$	$R(\mu) = N\bar{y}^2 = SSM$	[LM 246]
$E(y_{ij}) = \mu + \alpha_i$	$R(\mu,\alpha) = \sum_{i=1}^{a} n_i \bar{y}_{i.}^2$	[LM 246]
$E(y_{ijk}) = \mu + \alpha_i + \beta_j$	$R(\mu,\alpha,\beta)$: difficult	[(63), LM 293 (69), LM 297]
$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$	$R(\mu,\alpha,\beta,\gamma) = \sum_{i=1}^{a} \sum_{j=1}^{b} n_{ij} \bar{y}_{ij.}^2$	[(61), LM 292]

These are considered in detail later in these notes.

CHAPTER 4

4. TWO ELEMENTARY MODELS

This chapter of notes follows Chapter 6 of the text very closely and for many topics refers the reader directly to the text. It considers two models: that for the 1-way classification,

$$y_{ij} = \mu + \alpha_i + e_{ij} ,$$

which is dealt with here in detail (pages 52-55), and that for the 2-way nested (hierarchical) classification,

$$y_{ijk} = \mu + \alpha_i + \beta_{ij} + e_{ij} ,$$

which is dealt with only in outline (pages 55-57).

4.1. The 1-way classification

a. Example Table 6.1, LM 229

Class	Observations	Total	Mean
1	74, 68, 77	219	73
2	76, 80	156	78
3	85, 93	178	89
		553	

b. Model $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$

$$\begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{21} \\ y_{22} \\ y_{31} \\ y_{32} \end{bmatrix} = \begin{bmatrix} 74 \\ 68 \\ 77 \\ 76 \\ 80 \\ 85 \\ 93 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{21} \\ e_{22} \\ e_{31} \\ e_{32} \end{bmatrix} . \qquad (24), \text{ LM } 230$$

c. Normal equations

$$\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^{o} = \underset{\sim}{X}'\underset{\sim}{y} \qquad (2),\ \text{LM } 227$$

$$\begin{bmatrix} 7 & 3 & 2 & 2 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} \begin{bmatrix} \mu^{o} \\ \alpha_1^{o} \\ \alpha_2^{o} \\ \alpha_3^{o} \end{bmatrix} = \begin{bmatrix} 553 \\ 219 \\ 156 \\ 178 \end{bmatrix} \qquad (30),\ \text{LM } 232$$

$$\underset{\sim}{b}^{o} = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{y} \qquad (3),\ \text{LM } 227$$

$$\underset{\sim}{G} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \qquad (32) \text{ and } (35),\ \text{LM } 233$$

$$\underset{\sim}{b}^{o} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 553 \\ 219 \\ 156 \\ 178 \end{bmatrix} = \begin{bmatrix} 0 \\ 73 \\ 78 \\ 89 \end{bmatrix} \qquad (34),\ \text{LM } 233$$

$$\underset{\sim}{H} = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X} \qquad (33),\ \text{LM } 233$$

$$= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 7 & 3 & 2 & 2 \\ 3 & 3 & 0 & 0 \\ 2 & 0 & 2 & 0 \\ 2 & 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \qquad (36),\ \text{LM } 233$$

d. Analysis of variance   Table 6.2, and preceding equations, LM 234

Source of Variation	d.f.	Sum of Squares		
Mean	1	$R(\mu)$	$= N\bar{y}_{..}^2$	$= 43,687$
Classes	$a-1 = 2$	$R(\alpha\|\mu)$	$= R(\mu,\alpha) - R(\mu)$	
		i.e., SSR_m	$= \sum_{i=1}^{a} y_{i.}^2/n_i - N\bar{y}_{..}^2$	
			$= 43,997 - 43,687$	$= 310$
Error	$N-a = 4$	SSE	$= SST - R(\mu,\alpha)$	
			$= \sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2 - \sum_{i=1}^{a} y_{i.}^2/n_i$	
			$= 44,079 - 43,997$	$= 82$
Total	$N = 7$	SST	$= \Sigma\Sigma y_{ij}^2$	$= 44,079$

e. Estimable functions   LM 235-238

Basic estimable function: $\mu + \alpha_i$

b.l.u.e. of $\mu + \alpha_i$ is: $\widehat{\mu + \alpha_i} = \mu^o + \alpha_i^o = 0 + \bar{y}_{i.} = \bar{y}_{i.}$ (42)

$$v(\widehat{\mu + \alpha_i}) = \sigma^2/n_i \tag{44}$$

b.l.u.e. of $\Sigma\lambda_i(\mu + \alpha_i)$ is: $\widehat{\sum_{i=1}^{a} \lambda_i(\mu + \alpha_i)} = \sum_{i=1}^{a} \lambda_i \bar{y}_{i.}$ (45),(46)

$$v\left[\widehat{\sum_{i=1}^{a} \lambda_i(\mu + \alpha_i)}\right] = \sum_{i=1}^{a} (\lambda_i^2/n_i)\sigma^2 \tag{47}$$

Note: (i) μ is not estimable [LM 237]

(ii) α_i is not estimable [LM 237]

(iii) $(\Sigma\lambda_i)\mu + \Sigma\lambda_i\alpha_i$ is estimable for any λ_i's: [LM 237]

$$\mu + \sum_{i=1}^{a} n_i\alpha_i/N \quad \text{has b.l.u.e.} \quad \bar{y}_{..} \tag{52}$$

$$\mu + \sum_{i=1}^{a} \alpha_i/a \quad \text{has b.l.u.e.} \quad \Big(\sum_{i=1}^{a} \bar{y}_{i.}\Big)/a \tag{53}$$

(iv) $\Sigma\lambda_i\alpha_i$ is estimable for $\Sigma\lambda_i = 0$, with b.l.u.e. $\Sigma\lambda_i\bar{y}_{i.}$

(v) $\alpha_i - \alpha_k$ is estimable

$$\widehat{\alpha_i - \alpha_k} = \bar{y}_{i.} - \bar{y}_{k.} \tag{55}$$

$$v(\widehat{\alpha_i - \alpha_k}) = \sigma^2(1/n_i + 1/n_k)$$

f. Hypothesis testing

$$H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m} \qquad F(H) = \frac{(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})}{s\hat{\sigma}^2}$$

Example: LM 239

g. Explaining F-statistics of the ANOVA table

$F(M) = \dfrac{N\bar{y}^2}{\hat{\sigma}^2}$ tests $H: N\mu + \Sigma n_i\alpha_i = 0$ [LM 239]

$F(R_m) = \dfrac{R(\alpha|\mu)}{(a-1)\hat{\sigma}^2}$ tests H: all α_i's equal. [LM 240]

h. Independent and orthogonal contrasts

Hypotheses consisting of contrasts of the form

$$H: \alpha_1 - \alpha_2 = 0$$
$$\alpha_1 + \alpha_2 - 2\alpha_3 = 0$$

can be tested. Although these are said to be orthogonal in the usual sense of the word, their individual numerator sums of squares are not independent (because of the unbalanced data and the fact that $\underset{\sim}{k}_i'\underset{\sim}{G}\underset{\sim}{k}_j \neq 0$). Orthogonal contrasts satisfying this requirement are, for example,

$$H: \alpha_1 - \alpha_2 \qquad = 0$$
$$3\alpha_1 + 2\alpha_2 - 5\alpha_3 = 0 .$$

i. Restricted models Table 6.3, LM 245

j. Balanced data LM 245

$$n_i = n \text{ for all } i = 1, \cdots, a$$

$$\Sigma\alpha_i^o = 0 \quad \text{provides} \quad \mu^o = \bar{y}_{..}$$
$$\alpha_i^o = \bar{y}_{i.} - \bar{y}_{..}$$

$$\Sigma\alpha_i = 0 \quad \text{provides} \quad \hat{\mu} = \bar{y}_{..}$$
$$\hat{\alpha}_i = \bar{y}_{i.} - \bar{y}_{..} .$$

4.2. The 2-way nested classification

a. Example Table 6.5, LM 249

b. Model $\underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e}$ (67), LM 250

c. Normal equations

$(\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^o = \underset{\sim}{X}'\underset{\sim}{y})$ (68) and (69), LM 251

$\underset{\sim}{b}^o = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{y}$ (71) and (72), LM 252

$\underset{\sim}{G}$ (73), LM 252

d. Analysis of variance LM 252-253

Source of Variation	d.f.	Sum of Squares		
Mean	1	$N\bar{y}^2_{\cdot\cdot\cdot}$		$= 432$
α-classes	$a-1$	$R(\alpha\|\mu)$	$= R(\mu,\alpha) - R(\mu)$	
			$= \sum_{i=1}^{a} y^2_{i\cdot\cdot}/n_{i\cdot} - N\bar{y}^2$	
			$= 456 - 432$	$= 24$
β-classes, within α-classes	$b_{\cdot} - a$	$R(\beta{:}\alpha\|\mu,\alpha)$	$= R(\mu,\alpha,\beta{:}\alpha)$	
			$= R(\mu,\alpha,\beta{:}\alpha) - R(\mu,\alpha)$	
			$= \sum_{i=1}^{a}\sum_{j=1}^{b_i} y^2_{ij\cdot}/n_{ij} - \Sigma y^2_{i\cdot\cdot}/n_{i\cdot}$	
			$= 516 - 456$	$= 60$
Error	$N - b_{\cdot}$	SSE	$= \underset{\sim}{y}'\underset{\sim}{y} - R(\mu,\alpha,\beta{:}\alpha)$	
			$= \sum_{i}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} y^2_{ijk} - \sum_{i=1}^{a}\sum_{j=1}^{b_i} y^2_{ij\cdot}/n_{ij}$	
			$= 542 - 516$	$= 26$
Total	N	SST	$= \underset{\sim}{y}'\underset{\sim}{y} = \sum_{i=1}^{a}\sum_{j=1}^{b_i}\sum_{k=1}^{n_{ij}} y^2_{ijk}$	$= 542$

e. Estimable functions Table 6.8, LM 254

Basic estimable function: $\mu + \alpha_i + \beta_{ij}$

$$\widehat{\mu + \alpha_i + \beta_{ij}} = \bar{y}_{ij\cdot}$$

$$\widehat{\beta_{ij} - \beta_{ij'}} = \bar{y}_{ij\cdot} - \bar{y}_{ij'\cdot} \quad \text{for } j \neq j'$$

Note: μ, α_i, $\alpha_i - \alpha_k$ are not estimable.

f. Explaining F-statistics of the ANOVA table LM 255-256

$F(M)$ tests $H: N\mu + \sum_{i=1}^{a} n_i\alpha_i + \sum_{i=1}^{a}\sum_{j=1}^{b_i} n_{ij}\beta_{ij} = 0$.

$F(\alpha|\mu)$ tests $H: \alpha_i + \sum_{j=1}^{b_i} n_{ij}\beta_{ij}/n_{i.} = \alpha_{i'} + \sum_{j=1}^{b_{i'}} n_{i'j}\beta_{i'j}/n_{i'.}$

for all $i \neq i'$,

or, equivalently, $H: \alpha_i + \sum_{j=1}^{b_i} n_{ij}\beta_{ij}/n_{i.}$ equal for all $i = 1, \cdots, a$.

$F(\beta{:}\alpha|\mu,\alpha)$ tests $H: \beta_{ij} = \beta_{ij'}$ for all i, and all $j \neq j'$.

g. Restricted models LM 256-257

If $\sum_{j=1}^{b_i} n_{ij}\beta_{ij} = 0$ for all $i = 1, \cdots, a$ is part of the model, then

$F(\alpha|\mu)$ tests $H: \alpha_i$ all equal.

If $\sum_{i=1}^{a} n_i\alpha_i = 0$ and $\sum_{j=1}^{b_i} n_{ij}\beta_{ij} = 0$ for all $i = 1, \cdots, a$ are both part of the model, then

$F(\mu)$ tests $H: \mu = 0$.

h. Balanced data LM 257

$n_{ij} = n$ for all i and j, and $b_i = b$ for all i

$\sum_{i=1}^{a} \alpha_i^o = 0, \quad \sum_{j=1}^{b} \beta_{ij}^o = 0$ provide

$$\mu^o = \bar{y}_{...}$$
$$\alpha_i^o = \bar{y}_{i..} - \bar{y}_{...}$$
$$\beta_{ij}^o = \bar{y}_{ij.} - \bar{y}_{i..}$$

$\sum_{i=1}^{a} \alpha_i = 0, \quad \sum_{j=1}^{b} \beta_{ij} = 0$ provide

$$\hat{\mu} = \bar{y}_{...}$$
$$\hat{\alpha}_i = \bar{y}_{i..} - \bar{y}_{...}$$
$$\hat{\beta}_{ij} = \bar{y}_{ij.} - \bar{y}_{i..} \ .$$

CHAPTER 5

5. THE 2-WAY CROSSED CLASSIFICATION WITHOUT INTERACTION

This chapter covers Section 7.1, LM 261-286. It deals with the model

$$y_{ij} = \mu + \alpha_i + \beta_j + e_{ij} \qquad \begin{array}{r} i = 1, \cdots, a \\ j = 1, \cdots, b \\ n_{ij} = 0 \text{ or } 1\ . \end{array}$$

Example Table 7.1, LM 262

Table 7.1. Number of Seconds (Beyond 3 Minutes) Taken to Boil 2 Quarts of Water

Brand of Stove	Make of Pan A	B	C	Total	No. of Observations	Mean
X	18	12	24	54	(3)	18
Y	—	—	9	9	(1)	9
Z	3	—	15	18	(2)	9
W	6	3	18	27	(3)	9
Total	27	15	66	108		
No. of Observations	(3)	(2)	(4)		(9)	
Mean	9	$7\frac{1}{2}$	$16\frac{1}{2}$			12

Model LM 263 $\underset{\sim}{y} = \underset{\sim}{Xb} + \underset{\sim}{e}$

$$\begin{bmatrix} 18 \\ 12 \\ 24 \\ 9 \\ 3 \\ 15 \\ 6 \\ 3 \\ 18 \end{bmatrix} = \begin{bmatrix} y_{11} \\ y_{12} \\ y_{13} \\ y_{23} \\ y_{31} \\ y_{33} \\ y_{41} \\ y_{42} \\ y_{43} \end{bmatrix} = \left[\begin{array}{c|cccc|ccc} 1 & 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & 1 & \cdot \\ 1 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & 1 & \cdot & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & 1 & \cdot & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & 1 & \cdot & \cdot & \cdot & 1 \\ 1 & \cdot & \cdot & \cdot & 1 & 1 & \cdot & \cdot \\ 1 & \cdot & \cdot & \cdot & 1 & \cdot & 1 & \cdot \\ 1 & \cdot & \cdot & \cdot & 1 & \cdot & \cdot & 1 \end{array}\right] \left[\begin{array}{c} \mu \\ \hline \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \hline \beta_1 \\ \beta_2 \\ \beta_3 \end{array}\right] + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{23} \\ e_{31} \\ e_{33} \\ e_{41} \\ e_{42} \\ e_{43} \end{bmatrix} . \quad (2), \text{ LM 263}$$

Normal equations $\qquad \underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^{o} = \underset{\sim}{X}'\underset{\sim}{y}$

$$\left[\begin{array}{c|cccc|ccc} 9 & 3 & 1 & 2 & 3 & 3 & 2 & 4 \\ \hline 3 & 3 & . & . & . & 1 & 1 & 1 \\ 1 & . & 1 & . & . & . & . & 1 \\ 2 & . & . & 2 & . & 1 & . & 1 \\ 3 & . & . & . & 3 & 1 & 1 & 1 \\ \hline 3 & 1 & . & 1 & 1 & 3 & . & . \\ 2 & 1 & . & . & 1 & . & 2 & . \\ 4 & 1 & 1 & 1 & 1 & . & . & 4 \end{array}\right] \left[\begin{array}{c} \mu^o \\ \hline \alpha_1^o \\ \alpha_2^o \\ \alpha_3^o \\ \alpha_4^o \\ \hline \beta_1^o \\ \beta_2^o \\ \beta_3^o \end{array}\right] = \left[\begin{array}{c} y_{..} \\ \hline y_{1.} \\ y_{2.} \\ y_{3.} \\ y_{4.} \\ \hline y_{.1} \\ y_{.2} \\ y_{.3} \end{array}\right] = \left[\begin{array}{c} 108 \\ \hline 54 \\ 9 \\ 18 \\ 27 \\ \hline 27 \\ 15 \\ 66 \end{array}\right] . \qquad (3),\ \text{LM } 264$$

Solving the normal equations

There are no simple expressions for the elements of the solution vector $\underset{\sim}{b}^{o}$.

Example: LM 264-266 $\qquad$ General case: LM 266-267

Equations (14) and (16) - (18) are salient:

$$\alpha_i^o = \bar{y}_{i.} - \frac{1}{n_{i.}} \sum_{j=1}^{b-1} n_{ij}\beta_j^o \quad \text{for} \quad i = 1, 2, \cdots, a , \qquad (14)$$

$$\underset{\sim}{C}\underset{\sim}{\beta}_{b-1}^o = \underset{\sim}{r} \quad \text{with solution} \quad \underset{\sim}{\beta}_{b-1}^o = \underset{\sim}{C}^{-1}\underset{\sim}{r} \qquad (16)$$

where $\quad \underset{\sim}{C} = \{c_{jj'}\}$ and $\underset{\sim}{r} = \{r_j\}$ for $j = 1, \cdots, b-1$ $\qquad$ LM 267

with $$c_{jj} = n_{.j} - \sum_{i=1}^{a} \frac{n_{ij}^2}{n_{i.}} , \quad c_{jj'} = -\sum_{i=1}^{a} \frac{n_{ij}n_{ij'}}{n_{i.}} \quad \text{for} \quad j \neq j' \qquad (17)$$

and $$r_j = y_{.j} - \sum_{i=1}^{a} n_{ij}\bar{y}_{i.} \quad \text{for} \quad j = 1, \cdots, b-1 . \qquad (18)$$

A check on these calculations is provided by also calculating c_{bb}, c_{jb} and r_b and confirming that

$$\sum_{j'=1}^{b} c_{jj'} = 0 \quad \text{for all } j, \quad \text{and} \quad \sum_{j=1}^{b} r_j = 0 .$$

The solution $\underset{\sim}{\beta}_{b-1}^o$ in (16) is subscripted to emphasize that it has b - 1 and not b elements.

Matrix form of general case: LM 267-269.

Analysis of variance

$$R(\mu) = N\bar{y}^2_{..} \qquad R(\mu,\alpha) = \sum_{i=1}^{a} n_{i.}\bar{y}^2_{i.} \qquad R(\mu,\beta) = \sum_{j=1}^{b} n_{.j}\bar{y}^2_{.j}$$ LM 270

$$R(\beta|\mu,\alpha) = \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r}$$ (32), LM 271

$$\left.\begin{array}{ll} R(\alpha|\mu) & = R(\mu,\alpha) - R(\mu) \\ R(\beta|\mu,\alpha) & = \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r} \end{array}\right\} R(\mu,\alpha,\beta) - R(\mu) \left\{\begin{array}{ll} R(\beta|\mu) & = R(\mu,\beta) - R(\mu) \\ R(\alpha|\mu,\beta) & = R(\mu,\alpha) + \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r} - R(\mu,\beta) \end{array}\right.$$

TABLE 7.3. ANALYSES OF VARIANCE FOR 2-WAY CLASSIFICATION, NO INTERACTION

Source of Variation	Degrees of Freedom[1]	Sum of Squares[2]		Equation
		Table 7.3a: For fitting μ, and α and β after μ		
Mean, μ	1	$R(\mu)$	$= n_{..}\bar{y}^2_{..}$	(24)
α and β after μ	$a+b-2$	$R(\alpha,\beta\,|\,\mu)$	$= \sum_i n_{i.}\bar{y}^2_{i.} + \mathbf{r}'\mathbf{C}^{-1}\mathbf{r} - n_{..}\bar{y}^2_{..}$	(30)
Residual error[3]	N'	SSE	$= \sum_i\sum_j y^2_{ij} - \sum_i n_{i.}\bar{y}^2_{i.} - \mathbf{r}'\mathbf{C}^{-1}\mathbf{r}$	
Total	N	SST	$= \sum_i\sum_j y^2_{ij}$	
		Table 7.3b: For fitting μ, α after μ, and β after μ and α		
Mean, μ	1	$R(\mu)$	$= n_{..}\bar{y}^2_{..}$	(24)
α after μ	$a-1$	$R(\alpha\,|\,\mu)$	$= \sum_i n_{i.}\bar{y}^2_{i.} - n_{..}\bar{y}^2_{..}$	(31)
β after μ and α	$b-1$	$R(\beta\,|\,\mu,\alpha)$	$= \mathbf{r}'\mathbf{C}^{-1}\mathbf{r}$	(32)
Residual error	N'	SSE	$= \sum_i\sum_j y^2_{ij} - \sum_i n_{i.}\bar{y}^2_{i.} - \mathbf{r}'\mathbf{C}^{-1}\mathbf{r}$	
Total	N	SST	$= \sum_i\sum_j y^2_{ij}$	
		Table 7.3c: For fitting μ, β after μ, and α after μ and β		
Mean, μ	1	$R(\mu)$	$= n_{..}\bar{y}^2_{..}$	(24)
β after μ	$b-1$	$R(\beta\,|\,\mu)$	$= \sum_j n_{.j}\bar{y}^2_{.j} - n_{..}\bar{y}^2_{..}$	(37)
α after β and μ	$a-1$	$R(\alpha\,|\,\mu,\beta)$	$= \sum_i n_{i.}\bar{y}^2_{i.} + \mathbf{r}'\mathbf{C}^{-1}\mathbf{r} - \sum_j n_{.j}\bar{y}^2_{.j}$	(39)
Residual error	N'	SSE	$= \sum_i\sum_j y^2_{ij} - \sum_i n_{i.}\bar{y}^2_{i.} - \mathbf{r}'\mathbf{C}^{-1}\mathbf{r}$	
Total	N	SST	$= \sum_i\sum_j y^2_{ij}$	

LM 275

[1] $N \equiv n_{..}$ and $N' = N - a - b + 1$.
[2] $\mathbf{r}'\mathbf{C}^{-1}\mathbf{r}$ is obtained from equations (16)–(18).
[3] Summations are for $i = 1, 2, \ldots, a$ and $j = 1, 2, \ldots, b$.

Example: Table 7.2, LM 272 Conclusions: Table 7.4, LM 278

Estimable functions LM 279-280

Basic estimable function: $\mu + \alpha_i + \beta_j$ for $n_{ij} = 1$ (42)

$$\widehat{\mu + \alpha_i + \beta_j} = \mu^o + \alpha_i^o + \beta_j^o \qquad (43)$$

$$\widehat{\alpha_i - \alpha_h} = \alpha_i^o - \alpha_h^o \qquad (44)$$

$$\widehat{\beta_j - \beta_k} = \beta_j^o - \beta_k^o \;.$$

Consider $\underset{\sim}{G} = \underset{\sim}{G}' = \{g_{rs}\}$ $r, s = 1, 2, \cdots, a + b + 1$. If the diagonal elements of $\underset{\sim}{G}$ corresponding to α_i and α_h are g_{ii} and g_{hh}, respectively; and if g_{ih} is at the intersection of the corresponding row and column, then

$$v(\widehat{\alpha_i - \alpha_h}) = (g_{ii} + g_{hh} - 2g_{ih})\sigma^2 \;. \qquad (45)$$

A form for $\underset{\sim}{G}$ is given in (21), LM 268.

Explaining F-statistics of the ANOVA table

$F(M)$ tests $H: N\mu + \sum_{i=1}^{a} n_{i.}\alpha_i + \sum_{j=1}^{b} n_{.j}\beta_j = 0$.

$F(\alpha|\mu)$ tests $H: \alpha_i + \sum_{j=1}^{b} n_{ij}\beta_j/n_{i.}$ equal for all i . [LM 283]

$F(\beta|\mu,\alpha)$ tests $H: \beta_j$'s all equal.

$F(\beta|\mu)$ tests $H: \beta_j + \sum_{i=1}^{a} n_{ij}\alpha_i/n_{.j}$ equal for all j . [(48), LM 282]

$F(\alpha|\mu,\beta)$ tests $H: \alpha_i$'s all equal.

Examples: LM 281-283

<u>Balanced data</u> LM 284-285

$$n_{ij} = 1 \quad \text{for all } i \text{ and } j$$

$$R(\alpha|\mu) \equiv R(\alpha|\mu,\beta) \equiv \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{i.} - \bar{y}_{..})^2$$

$$R(\beta|\mu) \equiv R(\beta|\mu,\alpha) \equiv \sum_{i=1}^{a} \sum_{j=1}^{b} (\bar{y}_{.j} - \bar{y}_{..})^2$$

Analysis of variance: Table 7.5, LM 285

Estimable functions: $\widehat{\alpha_i - \alpha_h} = \bar{y}_{i.} - \bar{y}_{h.}$

$$\widehat{\beta_j - \beta_k} = \bar{y}_{.j} - \bar{y}_{.k}$$

CHAPTER 6

6. THE 2-WAY CROSSED CLASSIFICATION WITH INTERACTION

This chapter covers Section 7.2, LM 286-316. It deals with the model

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_{ij} + e_{ijk} \qquad \begin{array}{l} i = 1, \cdots, a \\ j = 1, \cdots, b \\ k = 1, \cdots, n_{ij} \\ n_{ij} \geq 0 \\ n_{ij} > 0 \text{ for s cells.} \end{array}$$

Example Tables 7.6 and 7.6a, LM 287

TABLE 7.6. WEIGHT[1] OF GRAIN (OUNCES) FROM $4' \times 4'$ TRIAL PLOTS

Treatment	Variety 1	Variety 2	Variety 3	Variety 4	Totals
1	8 13 9 30 (3) 10 [2] $y_{11\cdot}(n_{11})\bar{y}_{11\cdot}$		12 12 (1) 12 $y_{13\cdot}(n_{13})\bar{y}_{13\cdot}$	7 11 18 (2) 9 $y_{14\cdot}(n_{14})\bar{y}_{14\cdot}$	60 (6) 10 $y_{1\cdot\cdot}(n_{1\cdot})\bar{y}_{1\cdot\cdot}$
2	6 12 18 (2) 9 $y_{21\cdot}(n_{21})\bar{y}_{21\cdot}$	12 14 26 (2) 13 $y_{22\cdot}(n_{22})\bar{y}_{22\cdot}$			44 (4) 11 $y_{2\cdot\cdot}(n_{2\cdot})\bar{y}_{2\cdot\cdot}$
3		9 7 16 (2) 8 $y_{32\cdot}(n_{32})\bar{y}_{32\cdot}$	14 16 30 (2) 15 $y_{33\cdot}(n_{33})\bar{y}_{33\cdot}$	10 14 11 13 48 (4) 12 $y_{34\cdot}(n_{34})\bar{y}_{34\cdot}$	94 (8) 11¾ $y_{3\cdot\cdot}(n_{3\cdot})\bar{y}_{3\cdot\cdot}$
Totals	48 (5) 9.6 $y_{\cdot 1\cdot}(n_{\cdot 1})\bar{y}_{\cdot 1\cdot}$	42 (4) 10½ $y_{\cdot 2\cdot}(n_{\cdot 2})\bar{y}_{\cdot 2\cdot}$	42 (3) 14 $y_{\cdot 3\cdot}(n_{\cdot 3})\bar{y}_{\cdot 3\cdot}$	66 (6) 11 $y_{\cdot 4\cdot}(n_{\cdot 4})\bar{y}_{\cdot 4\cdot}$	198 (18) 11 $y_{\cdots}(n_{\cdot\cdot})\bar{y}_{\cdots}$

[1] The basic entries in the table are weights from individual plots.
[2] In each triplet of numbers the first is a total weight, the second (in parentheses) is the number of plots in the total and the third is the mean.

TABLE 7.6a. n_{ij}-VALUES OF TABLE 7.6

i	$j = 1$	$j = 2$	$j = 3$	$j = 4$	Totals: $n_{i.}$
1	3	-	1	2	6
2	2	2	-	-	4
3	-	2	2	4	8
Totals: $n_{.j}$	5	4	3	6	$n_{..} = 18$

Model

$$\begin{bmatrix}8\\13\\9\\12\\7\\11\\6\\12\\12\\14\\9\\7\\14\\16\\10\\14\\11\\13\end{bmatrix}=\begin{bmatrix}y_{111}\\y_{112}\\y_{113}\\y_{131}\\y_{141}\\y_{142}\\y_{211}\\y_{212}\\y_{221}\\y_{222}\\y_{321}\\y_{322}\\y_{331}\\y_{332}\\y_{341}\\y_{342}\\y_{343}\\y_{344}\end{bmatrix}=\begin{array}{c}\begin{array}{cccccccccccccccc}\mu&\alpha_1&\alpha_2&\alpha_3&\beta_1&\beta_2&\beta_3&\beta_4&\gamma_{11}&\gamma_{13}&\gamma_{14}&\gamma_{21}&\gamma_{22}&\gamma_{32}&\gamma_{33}&\gamma_{34}\end{array}\\ \begin{bmatrix}1&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot\\1&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot\\1&\cdot&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1&\cdot\\1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1\\1&\cdot&\cdot&1&\cdot&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&1\end{bmatrix}\end{array}\begin{bmatrix}\mu\\\alpha_1\\\alpha_2\\\alpha_3\\\beta_1\\\beta_2\\\beta_3\\\beta_4\\\gamma_{11}\\\gamma_{13}\\\gamma_{14}\\\gamma_{21}\\\gamma_{22}\\\gamma_{32}\\\gamma_{33}\\\gamma_{34}\end{bmatrix}+\begin{bmatrix}e_{111}\\e_{112}\\e_{113}\\e_{131}\\e_{141}\\e_{142}\\e_{211}\\e_{212}\\e_{221}\\e_{222}\\e_{321}\\e_{322}\\e_{331}\\e_{332}\\e_{341}\\e_{342}\\e_{343}\\e_{344}\end{bmatrix}. \qquad (52)\text{ , LM 289}$$

Normal equations

$$\begin{array}{c}\begin{array}{cccccccccccccccc}\mu^o&\alpha_1^o&\alpha_2^o&\alpha_3^o&\beta_1^o&\beta_2^o&\beta_3^o&\beta_4^o&\gamma_{11}^o&\gamma_{13}^o&\gamma_{14}^o&\gamma_{21}^o&\gamma_{22}^o&\gamma_{32}^o&\gamma_{33}^o&\gamma_{34}^o\end{array}\\ \begin{bmatrix}18&6&4&8&5&4&3&6&3&1&2&2&2&2&2&4\\6&6&\cdot&\cdot&3&\cdot&1&2&3&1&2&\cdot&\cdot&\cdot&\cdot&\cdot\\4&\cdot&4&\cdot&2&2&\cdot&\cdot&\cdot&\cdot&\cdot&2&2&\cdot&\cdot&\cdot\\8&\cdot&\cdot&8&\cdot&2&2&4&\cdot&\cdot&\cdot&\cdot&\cdot&2&2&4\\5&3&2&\cdot&5&\cdot&\cdot&\cdot&3&\cdot&\cdot&2&\cdot&\cdot&\cdot&\cdot\\4&\cdot&2&2&\cdot&4&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&2&2&\cdot&\cdot\\3&1&\cdot&2&\cdot&\cdot&3&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&2&\cdot\\6&2&\cdot&4&\cdot&\cdot&\cdot&6&\cdot&\cdot&2&\cdot&\cdot&\cdot&\cdot&4\\3&3&\cdot&\cdot&3&\cdot&\cdot&\cdot&3&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\1&1&\cdot&\cdot&\cdot&\cdot&1&\cdot&\cdot&1&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot\\2&2&\cdot&\cdot&\cdot&\cdot&\cdot&2&\cdot&\cdot&2&\cdot&\cdot&\cdot&\cdot&\cdot\\2&\cdot&2&\cdot&2&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&2&\cdot&\cdot&\cdot&\cdot\\2&\cdot&2&\cdot&\cdot&2&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&2&\cdot&\cdot&\cdot\\2&\cdot&\cdot&2&\cdot&2&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&2&\cdot&\cdot\\2&\cdot&\cdot&2&\cdot&\cdot&2&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&2&\cdot\\4&\cdot&\cdot&4&\cdot&\cdot&\cdot&4&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&\cdot&4\end{bmatrix}\end{array}\begin{bmatrix}\mu^o\\\alpha_1^o\\\alpha_2^o\\\alpha_3^o\\\beta_1^o\\\beta_2^o\\\beta_3^o\\\beta_4^o\\\gamma_{11}^o\\\gamma_{13}^o\\\gamma_{14}^o\\\gamma_{21}^o\\\gamma_{22}^o\\\gamma_{32}^o\\\gamma_{33}^o\\\gamma_{34}^o\end{bmatrix}=\begin{bmatrix}y_{..}\\y_{1..}\\y_{2..}\\y_{3..}\\y_{.1.}\\y_{.2.}\\y_{.3.}\\y_{.4.}\\y_{11.}\\y_{13.}\\y_{14.}\\y_{21.}\\y_{22.}\\y_{32.}\\y_{33.}\\y_{34.}\end{bmatrix}=\begin{bmatrix}198\\60\\44\\94\\48\\42\\42\\66\\30\\12\\18\\18\\26\\16\\30\\48\end{bmatrix} \qquad (53)\text{ , LM 290}$$

Solving the normal equations

$$\underset{\sim}{b}^{o'} = [\underset{\sim}{0}_{1\times(1+a+b)} \quad \underset{\sim}{\bar{y}}'] \qquad (55),\ \text{LM } 291$$

where $(\underset{\sim}{\bar{y}}')_{1\times s}$ = a vector of all $\bar{y}_{ij.}$'s for which $n_{ij} \neq 0$.

In our example

$$\underset{\sim}{b}^{o'} = [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \bar{y}_{11.}\ \bar{y}_{13.}\ \bar{y}_{14.}\ \bar{y}_{21.}\ \bar{y}_{22.}\ \bar{y}_{32.}\ \bar{y}_{33.}\ \bar{y}_{34.}] \qquad (56),\ \text{LM } 292$$

$$= [0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ \ 10\ \ 12\ \ 9\ \ 9\ \ 13\ \ 8\ \ 15\ \ 12\ ]$$

from Table 7.6.

The simplicity of this solution means that it is virtually unnecessary to derive the generalized inverse of $\underset{\sim}{X}'\underset{\sim}{X}$ that corresponds to $\underset{\sim}{b}^o$. That generalized inverse is, as is evident from (55) and the normal equations (53),

$$\underset{\sim}{G} = \begin{bmatrix} \underset{\sim}{0}_{(1+a+b)\times(1+a+b)} & \underset{\sim}{0}_{(1+a+b)\times s} \\ \underset{\sim}{0}_{s\times(1+a+b)} & \underset{\sim}{D}\{1/n_{ij}\} \end{bmatrix} \qquad (57),\ \text{LM } 292$$

where $\underset{\sim}{D}\{1/n_{ij}\}$ is a diagonal matrix of order s of the values $1/n_{ij}$ for the non-zero n_{ij}.

Analysis of variance - basic calculations, LM 292-3

$$R(\mu) = N\bar{y}^2_{...} \qquad R(\mu,\alpha) = \sum_{i=1}^{a} n_{i.}\bar{y}^2_{i..} \qquad R(\mu,\beta) = \sum_{j=1}^{b} n_{.j}\bar{y}^2_{.j.}$$

$$R(\mu,\alpha,\beta) = R(\mu,\alpha) + \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r} \qquad (63),\ \textit{et seq.},\ \text{LM } 293$$

and

$$R(\beta|\mu,\alpha) = \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r}$$

where $\underset{\sim}{C}$ is exactly the same as in the no-interaction model (except that instead of n_{ij} = 0 or 1, $n_{ij} \geq 0$); and $\underset{\sim}{r}$ is exactly the same, too, only with $y_{i..}$ and $y_{.j.}$ in place of $y_{i.}$ and $y_{.j}$. These are conceptually the same as in the no-interaction model. The additional term for the interaction model is

$$R(\mu,\alpha,\beta,\gamma) = \sum_{i=1}^{a} \sum_{j=1}^{b} y^2_{ij.}/n_{ij}\ , \qquad (61),\ \text{LM } 292$$

Analyses of Variance for the 2-way Crossed Classification

Row No.	Term	d.f.[1]	Sum of Squares: Mnemonic symbol	Description	Computational form[2]
1	Mean	1	$R(\mu)$	$= R(\mu)$	$= N\bar{y}^2_{...}$
2	Rows, adj. for μ	a-1	$R(\alpha\|\mu)$	$= R(\mu,\alpha) - R(\mu)$	$= \Sigma n_{i.}\bar{y}^2_{i..} - N\bar{y}^2_{...}$
3	Columns, adj. for μ and rows	b-1	$R(\beta\|\mu,\alpha)$	$= R(\mu,\alpha,\beta) - R(\mu,\alpha)$	$= \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r}$
4	Interaction, adj. for μ,α,β	s'	$R(\gamma\|\mu,\alpha,\beta)$	$= R(\mu,\alpha,\beta,\gamma) - R(\mu,\alpha,\beta)$	$= \Sigma\Sigma n_{ij}\bar{y}^2_{ij.} - \Sigma n_{i.}\bar{y}^2_{i..} + \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r}$
5	Error	N-s	SSE	$= \underset{\sim}{y}'\underset{\sim}{y} - R(\mu,\alpha,\beta,\gamma)$	$= \Sigma\Sigma\Sigma y^2_{ijk} - \Sigma\Sigma n_{ij}\bar{y}^2_{ij.} = \Sigma\Sigma\Sigma(y_{ijk} - \bar{y}_{ij.})^2$
6	Total	N	SST	$= \underset{\sim}{y}'\underset{\sim}{y}$	$= \Sigma\Sigma\Sigma y^2_{ijk}$

Alternate partitioning

Row No.	Term	d.f.	Mnemonic symbol	Description	Computational form
1a	Mean	1	$R(\mu)$	= same as line 1	
2a	Columns, adj. for μ	b-1	$R(\beta\|\mu)$	$= R(\mu,\beta) - R(\mu)$	$= \Sigma n_{.j}\bar{y}^2_{.j.} - N\bar{y}^2_{...}$
3a	Rows, adj. for μ and columns	a-1	$R(\alpha\|\mu,\beta)$	$= R(\mu,\alpha,\beta) - R(\mu,\beta)$	$= \Sigma n_{i.}\bar{y}^2_{i..} + \underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r} - \Sigma n_{.j}\bar{y}^2_{.j.}$
4a	Interaction, adj. for μ,α,β	s'	$R(\gamma\|\mu,\alpha,\beta)$	= same as line 4	
5a	Error	N-s	SSE	= same as line 5	
6a	Total	N	SST	= same as line 6	

1/ $N = n_{..}$, s = number of cells containing data, and $s' = s - a - b + 1$.

2/ $\underset{\sim}{r}'\underset{\sim}{C}^{-1}\underset{\sim}{r}$ is obtained as on page 61b of these notes.
Summations are for $i = 1, \cdots, a$, for $j = 1, \cdots, b$ and for $k = 1, \cdots, n_{ij}$, for $n_{ij} \neq 0$.

Example: Table 7.7, LM 294 Alternative expressions for general case: Table 7.8, LM 298-299.

and it, too, is easy. The difficult one is the one that is difficult in the no-interaction case:

$$R(\mu,\alpha,\beta)\begin{cases} : (63) - (65), \text{ LM } 293 \\ : (69) - (70), \text{ LM } 297 \end{cases}$$

Example: Table 7.7, LM 294

General case: Table 7.8, LM 298-299

Fitting main effects before interaction LM 300-301

$$R(\alpha|\mu) \quad \equiv R(\mu,\alpha) - R(\mu)$$

$$R(\alpha|\mu,\beta) \quad \equiv R(\mu,\alpha,\beta) - R(\mu,\beta)$$

In a formal way

$$R(\beta|\mu,\alpha,\gamma) \equiv R(\mu,\alpha,\beta,\gamma) - R(\mu,\alpha,\gamma). \tag{*}$$

But, in unrestricted models

$$R(\mu,\alpha,\beta,\gamma) = \text{SSR for fitting } E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

$$= \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij.}^2/n_{ij} \qquad [(61), \text{ LM } 292]$$

and

$$R(\mu,\alpha,\gamma) \quad = \text{SSR for fitting } E(y_{ijk}) = \mu + \alpha_i + \gamma_{ij}$$

$$= \text{SSR for fitting 2-way nested model}$$

$$= \sum_{i=1}^{a} \sum_{j=1}^{b} y_{ij.}^2/n_{ij}^2 \qquad [\text{LM } 252, \text{ 5 lines above Table } 6.6]$$

Hence $R(\beta|\mu,\alpha,\gamma) \equiv 0$.

With certain restricted models discussed in subsequent pages, a sum of squares that can be misleadingly given the symbol $R(\beta|\mu,\alpha,\gamma)$ can, in fact, be non-zero.

If in interpreting this symbol, the term $R(\mu,\alpha,\gamma)$ implicit therein, similar to (#) above, is interpreted as fitting $E(y_{ijk}) = \mu + \alpha_i + \gamma_{ij}$, then doing so poses the question: is it meaningful to think of γ_{ij} as an interaction, an interaction between α-effects and β-effects, with the β-effects themselves not being part of the model $E(y_{ijk}) = \mu + \alpha_i + \gamma_{ij}$?

Estimable functions LM 301-305

Basic estimable function: $\mu + \alpha_i + \beta_j + \gamma_{ij}$, for $n_{ij} > 0$

$$\widehat{\mu + \alpha_i + \beta_j + \gamma_{ij}} = \bar{y}_{ij.}$$

It is convenient to give a symbol to $\mu + \alpha_i + \beta_j + \gamma_{ij}$:

$$\mu_{ij} \equiv \mu + \alpha_i + \beta_j + \gamma_{ij} \qquad [(73),\ \text{LM } 302]$$

$$\hat{\mu}_{ij} = \bar{y}_{ij.}$$

and

$$v(\hat{\mu}_{ij}) = \sigma^2/n_{ij}$$

$\alpha_i - \alpha_{i'}$ is not estimable

$\beta_j - \beta_{j'}$ is not estimable

Due to the unbalanced nature of the data (unequal numbers of observations in the subclasses, including some empty cells), differences between α_i's are estimable only in the presence of corresponding β_j's and γ_{ij}'s.

Examples

[1] $$\alpha_i - \alpha_{i'} + \sum_{j=1}^{b} k_{ij}(\beta_j + \gamma_{ij}) - \sum_{j=1}^{b} k_{i'j}(\beta_j + \gamma_{i'j}) \qquad (76)$$

for $i \neq i'$, $\sum_{j=1}^{b} k_{ij} = 1 = \sum_{j=1}^{b} k_{i'j}$, $k_{ij} = 0$ when $n_{ij} = 0$.

This has b.l.u.e. $\sum_j k_{ij}\bar{y}_{ij.} - \sum_j k_{i'j}\bar{y}_{i'j.}$. (78)

[2] If there are m_i filled cells in row i:

$$\alpha_i - \alpha_{i'} + \sum_{\substack{j \text{ for} \\ n_{ij} \neq 0}} (\beta_j + \gamma_{ij})/m_i - \sum_{\substack{j \text{ for} \\ n_{i'j} \neq 0}} (\beta_j + \gamma_{i'j})/m_{i'}$$

has b.l.u.e. $\sum_{\substack{j \text{ for} \\ n_{ij} \neq 0}} \bar{y}_{ij.}/m_i - \sum_{\substack{j \text{ for} \\ n_{i'j} \neq 0}} \bar{y}_{i'j.}/m_{i'}$. [LM 303]

[3] For $k_{ij} = n_{ij}/n_{i.}$

$$\alpha_i - \alpha_{i'} + \sum_{j=1}^{b} n_{ij}(\beta_j + \gamma_{ij})/n_{i.} - \sum_{j=1}^{b} n_{i'j}(\beta_j + \gamma_{i'j})/n_{i'.} \qquad (84)$$

has b.l.u.e. $\bar{y}_{i..} - \bar{y}_{i'..}$

Provided cells (i,j), (i',j), (i,j') and (i',j') have data

$$\theta_{ij,i'j'} = \gamma_{ij} - \gamma_{i'j} - \gamma_{ij'} + \gamma_{i'j'} \qquad (82)$$

$$= \mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'}$$

is estimable, with b.l.u.e.

$$\hat{\theta}_{ij,i'j'} = \bar{y}_{ij.} - \bar{y}_{i'j.} - \bar{y}_{ij'.} + \bar{y}_{i'j'.} \qquad (83)$$

<u>Explaining F-statistics of the ANOVA table</u>

$F(M)$ tests $H: N\mu + \Sigma n_{i.}\alpha_i + \Sigma n_{.j}\beta_j + \Sigma\Sigma n_{ij}\gamma_{ij} = 0$ (97)

$F(\alpha|\mu)$ tests $H: \alpha_i + \sum_{j=1}^{b} n_{ij}(\beta_j + \gamma_{ij})/n_{i.}$

equal for all i [(100), LM 307]

$F(\beta|\mu)$ tests $H: \beta_j + \sum_{i=1}^{a} n_{ij}(\alpha_i + \gamma_{ij})/n_{.j}$

equal for all j [LM 308]

$$F(\alpha|\mu,\beta) \quad \text{tests} \quad H: (n_{i.} - \sum_{j=1}^{b} \frac{n_{ij}^2}{n_{.j}})\alpha_i - \sum_{i'\neq i}^{a} (\sum_{j=1}^{b} \frac{n_{ij}n_{i'j}}{n_{.j}})\alpha_{i'}$$

$$+ \sum_{j=1}^{b} (n_{ij} - \frac{n_{ij}^2}{n_{.j}})\gamma_{ij} - \sum_{i'\neq i}^{a} (\sum_{j=1}^{b} \frac{n_{ij}n_{i'j}}{n_{.j}})\gamma_{i'j} = 0$$

$$\text{for } i = 1, 2, \cdots, a-1 \qquad [(107), \text{ LM } 310]$$

$$F(\beta|\mu,\alpha) \quad \text{tests} \quad H: (n_{.j} - \sum_{i=1}^{a} \frac{n_{ij}^2}{n_{i.}})\beta_j - \sum_{j'\neq j}^{b} (\sum_{i=1}^{a} \frac{n_{ij}n_{ij'}}{n_{i.}})\beta_{j'}$$

$$+ \sum_{i=1}^{a} (n_{ij} - \frac{n_{ij}^2}{n_{i.}})\gamma_{ij} - \sum_{j'\neq j}^{b} (\sum_{i=1}^{a} \frac{n_{ij}n_{ij'}}{n_{i.}})\gamma_{ij'} = 0$$

$$\text{for } j = 1, 2, \cdots, b-1 \qquad [(106), \text{ LM } 309]$$

$$F(\gamma|\mu,\alpha,\beta) \quad \text{tests} \quad H: \left\{ \begin{array}{c} \text{any } s - a - b + 1 \text{ linearly} \\ \text{independent functions of} \\ \theta_{ij,i'j'}\text{'s, where such} \\ \text{functions are estimable} \\ \text{or estimable sums or} \\ \text{differences of } \theta\text{'s.} \end{array} \right\} = 0 \qquad [(110), \text{ LM } 311]$$

Restricted models

If $\sum_{j=1}^{b} n_{ij}(\beta_j + \gamma_{ij}) = 0$ for all $i = 1, \cdots, a$ is taken as part of the model, then from (84), $\alpha_i - \alpha_{i'}$ is estimable

$$\widehat{\alpha_i - \alpha_{i'}} = \bar{y}_{i..} - \bar{y}_{i'..}$$

and from (100), $F(\alpha|\mu)$ tests $H: \alpha_i$ all equal.

Note: Such restrictions utilize the sample, the n_{ij}'s.

Balanced data LM 315-316

$$n_{ij} = n \quad \text{for all } i = 1, \cdots, a \quad \text{and} \quad j = 1, \cdots, b$$

$$R(\alpha|\mu) \equiv R(\alpha|\mu,\beta) \equiv bn \sum_{i=1}^{a} (\bar{y}_{i..} - \bar{y}_{...})^2$$

[(123), LM 315]

$$R(\beta|\mu) \equiv R(\beta|\mu,\alpha) \equiv an \sum_{j=1}^{b} (\bar{y}_{.j.} - \bar{y}_{...})^2$$

Analysis of variance: Table 7.9, LM 316

Estimable functions: $\widehat{\alpha_i - \alpha_{i'}} = \bar{y}_{i..} - \bar{y}_{i'..}$

$$\widehat{\beta_j - \beta_{j'}} = \bar{y}_{.j.} - \bar{y}_{.j'.}$$

Customarily include the "usual restrictions" in the model

$$\sum_{i=1}^{a} \alpha_i = 0 \qquad \sum_{i=1}^{a} \gamma_{ij} = 0 \;\; \forall j$$

$$\sum_{j=1}^{b} \beta_j = 0 \qquad \sum_{j=1}^{b} \gamma_{ij} = 0 \;\; \forall i$$

CHAPTER 7

7. SOME OTHER ANALYSES

7.1. All cells filled

Unbalanced data are sometimes such that even though there are unequal numbers of observations in the cells, there is at least one observation in every cell; i.e., $n_{ij} \geq 1$ for all i and j. This is the all-cells-filled case. Since the mean of cell (i,j) is

$$\bar{y}_{ij.} = \mu + \alpha_i + \beta_j + \gamma_{ij} + \bar{e}_{ij.} ,$$

is present for all i and j, all differences between row effects and between column effects, in the presence of averaged interactions, are estimable:

$$\widehat{\alpha_i - \alpha_{i'} + \bar{\gamma}_{i.} - \bar{\gamma}_{i'.}} = \frac{1}{b}\left(\sum_{j=1}^{b} \bar{y}_{ij.} - \sum_{j=1}^{b} \bar{y}_{i'j.}\right)$$

and

$$\widehat{\beta_j - \beta_{j'} + \bar{\gamma}_{.j} - \bar{\gamma}_{.j'}} = \frac{1}{a}\left(\sum_{i=1}^{a} \bar{y}_{ij.} - \sum_{i=1}^{a} \bar{y}_{ij'.}\right),$$

where, with conventional notation,

$$\bar{\gamma}_{i.} = \sum_{j=1}^{b} \gamma_{ij}/b \quad \text{and} \quad \bar{\gamma}_{.j} = \sum_{i=1}^{a} \gamma_{ij}/a .$$

It is because all cells have data in them that these means apply for all i and j. And therefore, if we include in the model the "usual restrictions" of the balanced data case, namely

$$\sum_{j=1}^{b} \gamma_{ij} = 0 \text{ for all } i \quad \text{and} \quad \sum_{i=1}^{a} \gamma_{ij} = 0 \text{ for all } j ,$$

then the differences between row effects and between column effects become estimable:

$$\widehat{\alpha_i - \alpha_{i'}} = \frac{1}{b}\left(\sum_{j=1}^{b} \bar{y}_{ij.} - \sum_{j=1}^{b} \bar{y}_{i'j.}\right)$$

and

$$\widehat{\beta_j - \beta_{j'}} = \frac{1}{a}\left(\sum_{i=1}^{a} \bar{y}_{ij.} - \sum_{i=1}^{a} \bar{y}_{ij'.}\right).$$

Under these conditions, the hypotheses

$$H : \alpha_i\text{'s all equal}$$

and

$$H : \beta_j\text{'s all equal}$$

can therefore be tested. The numerator sums of squares for the F-statistics for making these tests are, respectively, the sums of squares SSA_w and SSB_w in the Weighted Squares of Means Analysis, Table 8.18, LM 370.

Correction

In Table 8.18, LM 370, the interaction sum of squares should read

$$SSAB_w = R(\gamma|\mu,\alpha,\beta) \text{ of Table 7.8, LM 298.}$$

For Table 8.19, LM 372, its calculated value is

$$SSAB_w = R(\gamma|\mu,\alpha,\beta) = 549\tfrac{21}{137} .$$

Note: Sums of squares in Table 8.19 do not "add up"; i.e., they do not add up to SST; and they are not meant to. The table is just a summary of sums of squares available for testing hypotheses.

7.2. Comments on hypothesis testing

We have devoted considerable space to describing certain hypothesis tests. But this is not to say that hypothesis testing is necessarily and always a good thing to do. Maybe as a statistical tool it is not appropriate in some situations.

After all, it is based on our deciding, regardless of data, that some probability such as .05 is so small as to represent an unlikely event; and if, on the basis of some hypothesis, we find from a set of data that that kind of event appears to have occurred, then we say that the hypothesis is not acceptable – we reject it. (If we wanted to "cheat", and change the probability value on which such a decision is made, we could no doubt decide to not reject it!) Thus maybe hypothesis testing is not the be-all and end-all of statistics; indeed, interval estimation, such as confidence interval determination, is considered much more de rigeur by many statisticians today.

And even if hypothesis testing is appropriate to the situation at hand, these hypotheses we have dealt with may not be the best ones to consider in terms of the real problems of interest. Indeed, this will frequently be the case.

Consider what we have done. We have partitioned a total sum of squares SST in one way or another, into terms denoted as R( ), $R(\alpha|\mu)$, for example. Then we have said "what hypothesis does this test?". And, in the case of the 2-way classification with interaction, we have found that

$$F(\alpha|\mu) = R(\alpha|\mu)/(a - 1)\hat{\sigma}^2$$

tests

$$H : \alpha_i + \frac{1}{n_{i.}} \Sigma n_{ij}(\beta_j + \gamma_{ij}) \text{ equal for all } i .$$

[See (100), LM 307.]

Several features of this hypothesis and its origin deserve comment.

(i) This hypothesis is not a hypothesis about α_i's only; and is certainly not a hypothesis that α_i's are all equal [which one might be led to think that it is, from the symbol $R(\alpha|\mu)$ and a knowledge of balanced-data hypothesis testing].

(ii) If the terms β_j and γ_{ij} of the model mean anything at all, then this hypothesis is a hotchpotch of β's and γ's as well as involving α's.

(iii) Furthermore, the hypothesis is based on the sample: it involves the n_{ij}'s, the numbers of observations in the sample. This does not seem to be good logic, testing a hypothesis about a population formulated in terms of the amount of data available in the sample.

(iv) Even if the preceding deficiencies are ignored, or accepted, then the hypothesis does not make very much sense. One reasonable interpretation is that it is

$$H : \left\{ \begin{array}{l} \text{row effects, in the presence} \\ \text{of average column and inter-} \\ \text{action effects, averaged for} \\ \text{each row according to the} \\ \text{number of observations in} \\ \text{each cell} \end{array} \right\} \begin{array}{c} \text{are} \\ \text{all} \\ \text{equal} \end{array} .$$

Why, then, have we taken such pains to show what is the consequence of using $R(\alpha|\mu)$ as a numerator sum of squares in an F-statistic? We have done so, not because the hypothesis is useful, but precisely for the opposite reason: because it is <u>not</u> useful. And many users of these sums of squares are not aware of this. They do not realize that the F-statistic based on $R(\alpha|\mu)$ is not testing $H : \alpha$'s all equal. It is therefore important that the true meaning, even though not useful, be understood.

Notice one other thing. The whole process is back-to-front so far as the logic of hypothesis testing is concerned: we have calculated sums of squares and ascertained what hypotheses they are testing. This is fine for giving us an understanding of what is being tested when these sums of squares are used in F-statistics. But the correct logic of hypothesis testing is just the opposite of this: first formulate a hypothesis, as

$$H : \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m} ,$$

where $\underset{\sim}{K}'\underset{\sim}{b}$ is estimable, i.e., $\underset{\sim}{K}' = \underset{\sim}{T}'\underset{\sim}{X}$ for some $\underset{\sim}{T}'$, and where $\underset{\sim}{K}'$ has full row rank;

and then test it with the F-statistic

$$F(H) = \frac{(\underset{\sim}{K}'\underset{\sim}{b}^{\circ} - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^{\circ} - \underset{\sim}{m})}{\hat{\sigma}^2 r(\underset{\sim}{K})}$$

where

$$\underset{\sim}{b}^{\circ} = \underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{y} \quad \text{and} \quad \underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{G}\underset{\sim}{X}'\underset{\sim}{X} = \underset{\sim}{X}'\underset{\sim}{X} .$$

It is true that not all hypotheses of interest can be tested (because such hypotheses are not always testable — i.e., cannot be expressed in terms of estimable functions). But for those that can be tested, the proper sequence logic is: formulate the hypothesis, and then calculate the corresponding F. Do not calculate an F and then wonder (or guess) what the hypothesis is.

Testing testable hypotheses can be done for as many hypotheses as one is interested in. True, the resulting F-statistics may not be statistically independent, but this is not uncommon. In fact, it is usually the case with the F-statistics in any analysis of variance table — even for balanced data. The numerator sums of squares of two F's may be independent, but the F's themselves are often not, simply because their denominators frequently use the same value of $\hat{\sigma}^2$. Lack of independence among the F-statistics is not, therefore, an uncommon occurrence.

7.3. Three-way and higher-order classifications   LM 332-340

All the difficulties evident in the 2-way crossed classification are simply aggravated in 3-way and higher-order classifications:

Large numbers of interactions

- and insufficient data to estimate them

Interactions of high order

- and no ability to interpret them
- and no data on some of them

Main-effects-only models

- avoid problems with interactions
- sequencing the fitting of factors
- n! sequences, see Table 8.2, LM 336
- R(a factor|all others) is useful

7.4. Covariance

General linear model theory can also be applied to covariance models. This yields very general computing procedures that are applicable to unbalanced data generally, and to covariance models that have not otherwise received much attention.

Summary of estimation: LM 347-348

Analysis of variance: LM 344-345

Some new models: Paper BU-342

CHAPTER 8

8. ALTERNATIVE MODELS AND COMPUTING PROCEDURES

8.1. μ_{ij}-models

The basic data for a 2-way crossed classification model are cell means $\bar{y}_{ij.}$ based on n_{ij} observations. The larger the n_{ij} are, the more information we have about the within-cell variation. But, insofar as information on effects of the factors on the y-variable is concerned, the basic data are just the cell means. In fact, for the (i,j)'th cell containing data, those data constitute a random sample from a population having a mean that we shall specify as μ_{ij} and variance σ^2; i.e.,

$$y_{ijk} = \mu_{ij} + e_{ijk} \tag{1}$$

with

$$E(e_{ijk}) = 0, \quad E(e^2_{ijk}) = \sigma^2 \quad \text{and} \quad \text{cov(any 2 different } e_{ijk}\text{'s)} = 0.$$

This is called the μ_{ij}-model. [LM 324]

Estimation and hypothesis testing in the μ_{ij}-model is very straightforward:

$$\begin{aligned} \hat{\mu}_{ij} &= \bar{y}_{ij.} \\ v(\hat{\mu}_{ij}) &= \sigma^2/n_{ij} \\ \text{cov(any 2 different } \hat{\mu}_{ij}\text{'s)} &= 0 . \end{aligned} \tag{2}$$

A variation of model (1) is one in which there may, as part of the model, be some restrictions on the μ_{ij}'s. This will be the case, for example, if we wish to use a model having no interactions between rows and columns. The absence of such interactions is defined by having

$$\mu_{ij} - \mu_{i'j} - \mu_{ij'} + \mu_{i'j'} = 0 \tag{3}$$

for $i \neq i'$ and $j \neq j'$, where cells (i,j), $(i'j)$, (i,j') and (i',j') all contain data. On defining $\boldsymbol{\mu} = \{\mu_{ij}\}$ as the vector of μ_{ij}'s corresponding to the cells containing data, restrictions such as (3) can be expressed as

$$\mathbf{P}\boldsymbol{\mu} = \mathbf{0} \tag{4}$$

for some known matrix $\mathbf{P}$. Estimation of $\boldsymbol{\mu}$ for the model consisting of both (1) and (4) is then in the form

$$\hat{\boldsymbol{\mu}}_r = \bar{\mathbf{y}} - \mathbf{DP}(\mathbf{P}'\mathbf{DP})^{-1}\mathbf{P}'\bar{\mathbf{y}} \tag{5}$$

where

$$\bar{\mathbf{y}} = \{\bar{y}_{ij.}\} \quad \text{for the cells containing data}$$

and

$$\mathbf{D} = \mathbf{D}\{1/n_{ij}\}$$

is the diagonal matrix of terms $1/n_{ij}$ corresponding to the terms in $\bar{\mathbf{y}}$. [Equation (5) is a special case of (97), LM 206.] Then, from (5)

$$v(\hat{\boldsymbol{\mu}}_r) = [\mathbf{D} - \mathbf{DP}(\mathbf{P}'\mathbf{DP})^{-1}\mathbf{P}'\mathbf{D}]\sigma^2 \ . \tag{6}$$

When equations such as (3) do not exist, then $\mathbf{P} \equiv \mathbf{0}$ and (5) and (6) reduce to (2).

Example

A simple case of 2 rows and 3 columns with just a single empty cell is the following:

✓	✓	✓
✓	✓	

where ✓ indicates the presence of data. In this case the restriction on the μ_{ij}'s is

$$\mu_{11} - \mu_{12} - \mu_{21} + \mu_{22} = 0 ,$$

for which

$$\underset{\sim}{P} = [1 \quad -1 \quad 0 \quad -1 \quad 1] ,$$

$$\hat{\underset{\sim}{\mu}}_r = \bar{\underset{\sim}{y}} - \underset{\sim\sim}{DP}(1/n_{11} + 1/n_{12} + 1/n_{21} + 1/n_{22})^{-1}(\bar{y}_{11\cdot} - \bar{y}_{12\cdot} - \bar{y}_{21\cdot} + \bar{y}_{22\cdot})$$

$$= \begin{bmatrix} \bar{y}_{11\cdot} - \lambda/n_{11} \\ \bar{y}_{12\cdot} + \lambda/n_{12} \\ \bar{y}_{13\cdot} \\ \bar{y}_{21\cdot} + \lambda/n_{21} \\ \bar{y}_{22\cdot} - \lambda/n_{22} \end{bmatrix} \quad \text{for} \quad \lambda = \frac{\bar{y}_{11\cdot} - \bar{y}_{12\cdot} - \bar{y}_{21\cdot} + \bar{y}_{22\cdot}}{\frac{1}{n_{11}} + \frac{1}{n_{12}} + \frac{1}{n_{21}} + \frac{1}{n_{22}}} .$$

The great utility of models like (1), or (1) and (4) together, is that all of them are of full rank and in all of them every μ_{ij} (corresponding to a cell containing data) is estimable; and so is every linear function of such μ_{ij}'s. This permits a researcher to test any linear hypotheses about those μ_{ij}'s that interests him. True it is, that these models do not have built-in definitions of what we familiarly call row and column effects. But they do enable a researcher to define row and column effects as any linear functions of cell means that seems appropriate in the face of empty cells. Thus in the example, to estimate an effect due to row 1, that effect might be defined as $\frac{1}{3}(\mu_{11} + \mu_{12} + \mu_{13})$; but to compare rows 1 and 2, it would probably be defined as $\frac{1}{2}(\mu_{11} + \mu_{12})$. There need be no confusion in having two such definitions: the first is the effect averaged over all columns, whereas the second is the effect averaged over only those columns wherein there are also data on row 2, so permitting a comparison of rows 1 and 2 over the same columns.

Estimation

The b.l.u.c. of $\Sigma k_{ij}\mu_{ij}$ is

$$\widehat{\Sigma k_{ij}\mu_{ij}} = \Sigma k_{ij}\hat{\mu}_{ij} = \underset{\sim}{k}'\hat{\underset{\sim}{\mu}} ,$$

on defining $\underset{\sim}{k}'$ as the row vector of k_{ij}'s. Then

$$v(\widehat{\Sigma k_{ij}\mu_{ij}}) = \underset{\sim}{k}'[\underset{\sim}{D} - \underset{\sim}{D}\underset{\sim}{P}(\underset{\sim}{P}'\underset{\sim}{D}\underset{\sim}{P})^{-1}\underset{\sim}{P}'\underset{\sim}{D}]\underset{\sim}{k}\sigma^2 .$$

An unbiased estimator of σ^2 is

$$\hat{\sigma}^2 = \frac{\sum\limits_{i}\sum\limits_{j}\sum\limits_{k}(y_{ijk} - \hat{\mu}_{ij})^2}{N - s + q}$$

where s is the number of cells containing data and q is the number of restrictions on the μ_{ij}'s (number of rows in $\underset{\sim}{P}$).

When there are no restrictions on the μ_{ij}'s, i.e., $\underset{\sim}{P} = \underset{\sim}{0}$ and $q = 0$, as is often the case, then

$$\widehat{\Sigma k_{ij}\mu_{ij}} = \Sigma k_{ij}\bar{y}_{ij.} , \quad \text{with variance} \quad \Sigma k_{ij}^2/n_{ij} ,$$

and

$$\hat{\sigma}^2 = \frac{\sum\limits_{i}\sum\limits_{j}\sum\limits_{k}(y_{ijk} - \bar{y}_{ij.})^2}{N - s} = SSE .$$

Hypothesis testing [LM 326 and 339]

The hypothesis

$$H : \Sigma k_{ij}\mu_{ij} = m$$

is tested by comparing

$$F(H) = \frac{(\Sigma k_{ij}\hat{\mu}_{ij} - m)^2}{\hat{\sigma}^2\underset{\sim}{k}'[\underset{\sim}{D} - \underset{\sim}{D}\underset{\sim}{P}(\underset{\sim}{P}'\underset{\sim}{D}\underset{\sim}{P})^{-1}\underset{\sim}{P}'\underset{\sim}{D}]\underset{\sim}{k}} \quad \text{against } F_{1,N-s+q} .$$

When $\underset{\sim}{P} = \underset{\sim}{0}$ the denominator is $\hat{\sigma}^2(\Sigma k_{ij}^2/n_{ij})$.

8.2. Some computational equivalences

(A skeletal view of paper BU-668-M, March 1979)

References

N = these notes
P = paper BU-668-M
LM = text

5 Methods of Calculating Sums of Squares

(i) Full rank, reparameterized models

(ii) "Indirect", invert-part-of-the inverse

(iii) $R(\cdot|\cdot)$ procedure

(iv) Weighted squares of means (all cells filled)

(v) Numerator of an F-statistic: $(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})$.

Example

7,9	6	2	24
8	4,8	12	32
24	18	14	56

Reference

P29

Model with Interaction

$$E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$$

Table A5: Partitionings of Total Sums of Squares

(a) Rows before Columns			(b) Columns before Rows		
Term	d.f.	Sum of Squares	Term	d.f.	Sum of Squares
$R(\mu)$	1	392	$R(\mu)$	1	392
$R(\alpha\|\mu)$	1	8	$R(\beta\|\mu)$	2	6
$R(\beta\|\mu,\alpha)$	2	$11\frac{7}{11}$	$R(\alpha\|\mu,\beta)$	1	$13\frac{7}{11}$
$R(\gamma\|\mu,\alpha,\beta)$	2	$36\frac{4}{11}$	$R(\gamma\|\mu,\alpha,\beta)$	2	$36\frac{4}{11}$
SSE	2	10	SSE	2	10
SST	8	458	SST	8	458

P29

Also, the sums of squares from the weighted squares of means analysis are

$SSA_w = 20$ and $SSB_w = 5\frac{1}{3}$. (A1), P29

The model equations for these data are

$$\begin{bmatrix} 7 \\ 9 \\ 6 \\ 2 \\ 8 \\ 4 \\ 8 \\ 12 \end{bmatrix} = \underset{\sim}{y} = \underset{\sim}{X}\underset{\sim}{b} + \underset{\sim}{e} = \left[\begin{array}{c|cc|ccc|cccccc} 1 & 1 & . & 1 & . & . & 1 & . & . & . & . & . \\ 1 & 1 & . & 1 & . & . & 1 & . & . & . & . & . \\ 1 & 1 & . & . & 1 & . & . & 1 & . & . & . & . \\ 1 & 1 & . & . & . & 1 & . & . & 1 & . & . & . \\ 1 & . & 1 & 1 & . & . & . & . & . & 1 & . & . \\ 1 & . & 1 & . & 1 & . & . & . & . & . & 1 & . \\ 1 & . & 1 & . & 1 & . & . & . & . & . & 1 & . \\ 1 & . & 1 & . & . & 1 & . & . & . & . & . & 1 \end{array}\right] \left[\begin{array}{c} \mu \\ \hline \alpha_1 \\ \alpha_2 \\ \hline \beta_1 \\ \beta_2 \\ \beta_3 \\ \hline \gamma_{11} \\ \gamma_{12} \\ \gamma_{13} \\ \gamma_{21} \\ \gamma_{22} \\ \gamma_{23} \end{array}\right] + \underset{\sim}{e}\,. \qquad \text{(A2), P30}$$

(Dots in a matrix represent zeros.)

The normal equations $\underset{\sim}{X}'\underset{\sim}{X}\underset{\sim}{b}^{\circ} = \underset{\sim}{X}'\underset{\sim}{y}$ resulting from this are

$$\left[\begin{array}{c|cc|ccc|cccccc} 8 & 4 & 4 & 3 & 3 & 2 & 2 & 1 & 1 & 1 & 2 & 1 \\ \hline 4 & 4 & . & 2 & 1 & 1 & 2 & 1 & 1 & . & . & . \\ 4 & . & 4 & 1 & 2 & 1 & . & . & . & 1 & 2 & 1 \\ \hline 3 & 2 & 1 & 3 & . & . & 2 & . & . & 1 & . & . \\ 3 & 1 & 2 & . & 3 & . & . & 1 & . & . & 2 & . \\ 2 & 1 & 1 & . & . & 2 & . & . & 1 & . & . & 1 \\ \hline 2 & 2 & . & 2 & . & . & 2 & . & . & . & . & . \\ 1 & 1 & . & . & 1 & . & . & 1 & . & . & . & . \\ 1 & 1 & . & . & . & 1 & . & . & 1 & . & . & . \\ 1 & . & 1 & 1 & . & . & . & . & . & 1 & . & . \\ 2 & . & 2 & . & 2 & . & . & . & . & . & 2 & . \\ 1 & . & 1 & . & . & 1 & . & . & . & . & . & 1 \end{array}\right] \left[\begin{array}{c} \mu^{\circ} \\ \hline \alpha_1^{\circ} \\ \alpha_2^{\circ} \\ \hline \beta_1^{\circ} \\ \beta_2^{\circ} \\ \beta_3^{\circ} \\ \hline \gamma_{11}^{\circ} \\ \gamma_{12}^{\circ} \\ \gamma_{13}^{\circ} \\ \gamma_{21}^{\circ} \\ \gamma_{22}^{\circ} \\ \gamma_{23}^{\circ} \end{array}\right] = \left[\begin{array}{c} 56 \\ \hline 24 \\ 32 \\ \hline 24 \\ 18 \\ 14 \\ \hline 16 \\ 6 \\ 2 \\ 8 \\ 12 \\ 12 \end{array}\right] . \qquad \text{(A3), P30}$$

Using a generalized inverse

$$(\underset{\sim}{X}'\underset{\sim}{X})^{-} = \text{diag}\{0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; \tfrac{1}{2} \;\; 1 \;\; 1 \;\; 1 \;\; \tfrac{1}{2} \;\; 1\}\,,$$

based on the general form given in equation (57), LM 292, a solution to (A3) is

$$\underset{\sim}{b}^{\circ\prime} = [0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 0 \;\; 8 \;\; 6 \;\; 2 \;\; 8 \;\; 6 \;\; 12]\,,$$

with

$$R(\mu,\alpha,\beta,\gamma) = \underset{\sim}{b}^{\circ\prime}\underset{\sim}{X}'\underset{\sim}{y} = [8(16) + 6(6) + 2(2) + 8(8) + 6(12) + 12(12)] = 448\,.$$

(i) Full rank, reparameterized model

Σ-restrictions

$$\begin{aligned} \dot\alpha_1 + \dot\alpha_2 &= 0 \Rightarrow \dot\alpha_2 = -\dot\alpha_1 \\ \dot\beta_1 + \dot\beta_2 + \dot\beta_3 &= 0 \Rightarrow \dot\beta_3 = -\dot\beta_1 - \dot\beta_2 \end{aligned} \qquad \text{(A10), P33}$$

$$\left.\begin{aligned} \dot\gamma_{11} + \dot\gamma_{12} + \dot\gamma_{13} &= 0 \\ \dot\gamma_{21} + \dot\gamma_{22} + \dot\gamma_{23} &= 0 \\ \dot\gamma_{11} + \dot\gamma_{21} &= 0 \\ \dot\gamma_{12} + \dot\gamma_{22} &= 0 \\ \dot\gamma_{13} + \dot\gamma_{23} &= 0 \end{aligned}\right\} \Rightarrow \begin{aligned} \dot\gamma_{11} &= \dot\gamma_{11} \\ \dot\gamma_{12} &= \dot\gamma_{12} \\ \dot\gamma_{13} &= -\dot\gamma_{11} - \dot\gamma_{12} \\ \dot\gamma_{21} &= -\dot\gamma_{11} \\ \dot\gamma_{22} &= -\dot\gamma_{12} \\ \dot\gamma_{23} &= \dot\gamma_{11} + \dot\gamma_{12} \end{aligned} \qquad \text{(A11), P33}$$

Model equations

$$\begin{bmatrix} 7 \\ 9 \\ 6 \\ 2 \\ 8 \\ 4 \\ 8 \\ 12 \end{bmatrix} = \left[\begin{array}{c|c|cc|cc} 1 & 1 & 1 & . & 1 & . \\ 1 & 1 & 1 & . & 1 & . \\ 1 & 1 & . & 1 & . & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 \\ 1 & -1 & 1 & . & -1 & . \\ 1 & -1 & . & 1 & . & -1 \\ 1 & -1 & . & 1 & . & -1 \\ 1 & -1 & -1 & -1 & 1 & 1 \end{array}\right] \begin{bmatrix} \dot\mu \\ \dot\alpha_1 \\ \dot\beta_1 \\ \dot\beta_2 \\ \dot\gamma_{11} \\ \dot\gamma_{12} \end{bmatrix} + \underset{\sim}{e} \, . \qquad \text{P34}$$

Normal equations

$$\left[\begin{array}{c|c|cc|cc} 8 & 0 & 1 & 1 & 1 & -1 \\ \hline 0 & 8 & 1 & -1 & 1 & 1 \\ \hline 1 & 1 & 5 & 2 & 1 & 0 \\ 1 & -1 & 2 & 5 & 0 & -1 \\ \hline 1 & 1 & 1 & 0 & 5 & 2 \\ -1 & 1 & 0 & -1 & 2 & 5 \end{array}\right] \left[\begin{array}{c} \hat{\dot\mu} \\ \hline \hat{\dot\alpha}_1 \\ \hline \hat{\dot\beta}_1 \\ \hat{\dot\beta}_2 \\ \hline \hat{\dot\gamma}_{11} \\ \hat{\dot\gamma}_{12} \end{array}\right] = \left[\begin{array}{c} 56 \\ \hline -8 \\ \hline 10 \\ 4 \\ \hline 18 \\ 4 \end{array}\right] . \qquad \text{(A12), P34}$$

(ii) "Indirect" Method: for full rank models

$$\underset{\sim}{y} = \underset{\sim}{X}_1\underset{\sim}{b}_1 + \underset{\sim}{X}_2\underset{\sim}{b}_2 + \underset{\sim}{e}$$

$$\begin{bmatrix} \hat{\underset{\sim}{b}}_1 \\ \hat{\underset{\sim}{b}}_2 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1} \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{y} \\ \underset{\sim}{X}_2'\underset{\sim}{y} \end{bmatrix} . \qquad (6),\ P4$$

Define $\begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{X}_1 & \underset{\sim}{X}_1'\underset{\sim}{X}_2 \\ \underset{\sim}{X}_2'\underset{\sim}{X}_1 & \underset{\sim}{X}_2'\underset{\sim}{X}_2 \end{bmatrix}^{-1} = \begin{bmatrix} \underset{\sim}{T}_{11} & \underset{\sim}{T}_{12} \\ \underset{\sim}{T}_{21} & \underset{\sim}{T}_{22} \end{bmatrix}, \Rightarrow \begin{bmatrix} \hat{\underset{\sim}{b}}_1 \\ \hat{\underset{\sim}{b}}_2 \end{bmatrix} = \begin{bmatrix} \underset{\sim}{T}_{11} & \underset{\sim}{T}_{12} \\ \underset{\sim}{T}_{21} & \underset{\sim}{T}_{22} \end{bmatrix} \begin{bmatrix} \underset{\sim}{X}_1'\underset{\sim}{y} \\ \underset{\sim}{X}_2'\underset{\sim}{y} \end{bmatrix} .$ (8), P4

Define $Q_{\underset{\sim}{b}_1} = \hat{\underset{\sim}{b}}_1'\underset{\sim}{T}_{11}^{-1}\hat{\underset{\sim}{b}}_1 .$ (107), LM 115

This is the "indirect" method of calculating sums of squares, sometimes also called "invert part of the inverse", because $\underset{\sim}{T}_{11}$ is part of the inverse in (8).

(iii) The $R(\cdot|\cdot)$ procedure [N50-1, P3-6, LM 246-7]

$$E(\underset{\sim}{y}) = \underset{\sim}{X}\underset{\sim}{b} \Rightarrow R(\underset{\sim}{b}) = \underset{\sim}{y}'\underset{\sim}{X}(\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{y} = \underset{\sim}{b}^{o\prime}\underset{\sim}{X}'\underset{\sim}{y} .$$

This can be called the R-algorithm:

$$R(\underset{\sim}{b}) = \underset{\sim}{b}^{o\prime}(\underset{\sim}{X}'\underset{\sim}{y}) = \Sigma\left(\begin{array}{c}\text{each element of solution vector}\\ \times\ \text{corresponding r.h.s. of normal equations}\end{array}\right) . \qquad P3$$

For $E(\underset{\sim}{y}) = \underset{\sim}{X}_1\underset{\sim}{b}_1 + \underset{\sim}{X}_2\underset{\sim}{b}_2$

$$R(\underset{\sim}{b}_1|\underset{\sim}{b}_2) = R(\underset{\sim}{b}_1, \underset{\sim}{b}_2) - R(\underset{\sim}{b}_1) .$$

(iv) Weighted squares of means [N67-8, P11, LM 369-72]

Only for the case of all cells filled.

Use $x_{ij} \equiv \bar{y}_{ij.}$, and for rows

$$SSA_w = \sum_{i=1}^{a} w_i(\bar{x}_{i.} - \bar{x}_{[1]})^2$$

where $w_i = b^2/\sum_j(1/n_{ij})$, and $\bar{x}_{[1]}$ is the weighted mean of the $\bar{x}_{i.}$'s using the w_i's. Columns are handled similarly.

(v) Numerator of an F-statistic [N43, P5, LM 190]

$H: \underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$, $\underset{\sim}{K}'\underset{\sim}{b}$ estimable, $\underset{\sim}{K}'$ full row rank.

$$F = Q/\hat{\sigma}^2 r(\underset{\sim}{K}') \quad \text{where} \quad r(\underset{\sim}{K}') \equiv \text{rank of } \underset{\sim}{K}' .$$

$$Q = (\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m}), \text{ for } \underset{\sim}{b}^o = (\underset{\sim}{X}'\underset{\sim}{X})^{-}\underset{\sim}{X}'\underset{\sim}{y} .$$

Relationships Between the Methods

Relationships depend upon model and data.

Full rank models

$$Q_{\underset{\sim}{b}_1} = R(\underset{\sim}{b}_1 | \underset{\sim}{b}_2) = Q \quad \text{for } H: \underset{\sim}{b}_1 = \underset{\sim}{0}\,. \qquad (9),\ P4$$

"indirect" numerator s.s.

For non-full rank models (e.g., 2-way crossed classification)

After making $E(y_{ijk}) = \mu + \alpha_i + \beta_j + \gamma_{ij}$ full rank by using Σ-restrictions,

$$Q_{\dot{\alpha}} = R^*(\underset{\sim}{\dot{\alpha}} | \dot{\mu}, \underset{\sim}{\dot{\beta}}, \underset{\sim}{\dot{\gamma}})_\Sigma = Q \quad \text{for } H: \dot{\alpha}_i\text{'s all zero.}$$

Further interpretation depends on model and data.

1. No interaction model (usually 0 or 1 observation)

$$Q_{\dot{\alpha}} = R(\alpha | \mu, \beta) = Q \quad \text{for } H: \alpha_i\text{'s all equal.}$$

(24), P10 (theory) and (A13), P35 (Example)

2. With interaction model (all cells filled)

$$Q_{\dot{\alpha}} = SSA_w = Q \quad \text{for } H: (\alpha_i + \bar{\gamma}_{i.})\text{'s all equal.}$$

(31), P12 (theory) and P36, top (Example)

3. With interaction (empty cells)

$$Q_{\dot{\alpha}} = Q \quad \text{for testing } H: \dot{\alpha}_i\text{'s all equal.}$$

The $\dot{\alpha}_i$'s, as functions of α_i's and other parameters of the over-parameterized model, depend upon the pattern of empty cells.

Example 1

✓	✓	✓
✓	✓	

✓ = data

$$\dot{\alpha}_1 = \tfrac{1}{2}(\alpha_1 - \alpha_2) + \tfrac{1}{4}(\gamma_{11} + \gamma_{12} - \gamma_{21} - \gamma_{22})\,. \qquad (33),\ P14$$

Example 2

✓	✓	✓
✓	✓	
✓		✓

$$\dot{\alpha}_1 = \alpha_1 - \tfrac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3) + \tfrac{1}{15}(4\gamma_{11} + 3\gamma_{12} + 3\gamma_{13} - 2\gamma_{21} - 3\gamma_{22} - 2\gamma_{31} - 3\gamma_{33})$$

(37), P16

$$\dot{\alpha}_2 = \alpha_2 - \tfrac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3) + \tfrac{1}{15}(-2\gamma_{11} - 4\gamma_{12} + \gamma_{13} + 6\gamma_{21} + 4\gamma_{22} - 4\gamma_{31} - \gamma_{33}).$$

Numerical Illustrations

Illustration is for terms concerning μ with data set

7,9	6	2	24
8	4,8	12	32
24	18	14	56

N77
P29
ACO's Data Set 3

First: $R(\mu) = N\bar{y}^2 = 8(56/8)^2 = 392.$ P29

Note: $R(\mu)$ does not depend on model, nor on the individual n_{ij}'s, nor on the pattern of empty cells (if any).

No Interaction Model: Σ-restrictions

$$Q_{\dot{\mu}} = R^*(\dot{\mu}|\dot{\alpha},\dot{\beta})_\Sigma = Q \quad \text{for } H : \mu + \tfrac{1}{2}(\alpha_1 + \alpha_2) + \tfrac{1}{3}(\beta_1 + \beta_2 + \beta_3) = 0 .$$

Normal equations are (A12) with $\dot{\gamma}$'s and $\dot{\gamma}$-equations removed: P34

$$\begin{bmatrix} 8 & 0 & 1 & 1 \\ 0 & 8 & 1 & -1 \\ 1 & 1 & 5 & 2 \\ 1 & -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \end{bmatrix} = \begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix}, \qquad (**)$$

$$\text{with solution} \quad \hat{\dot{\underset{\sim}{b}}} = \begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_1 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_2 \end{bmatrix} = \frac{1}{22(27)} \begin{bmatrix} 77 & 0 & -11 & -11 \\ 0 & 81 & -27 & 27 \\ -11 & -27 & 152 & -64 \\ -11 & 27 & -64 & 152 \end{bmatrix} \begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ -\frac{15}{11} \\ \frac{16}{11} \\ -\frac{16}{11} \end{bmatrix} .$$

(ii) "Indirect": $Q_{\dot{\mu}} = \hat{\dot{\mu}} T^{-1}_{\dot{\mu}\dot{\mu}} \hat{\dot{\mu}} = 7\left[\frac{77}{22(27)}\right]^{-1} 7 = 378 \neq R(\mu)$.

(iii) The $R(\cdot|\cdot)$ procedure

$$R^*(\dot\mu|\dot\alpha,\dot\beta)_\Sigma = R^*(\dot\mu,\dot\alpha,\dot\beta)_\Sigma - R^*(\dot\alpha,\dot\beta)_\Sigma \ .$$

Because the model with $\dot\mu$, $\dot\alpha$ and $\dot\beta$ is just a reparameterization of the one with μ, α and β, which is the full model being dealt with,

$$\begin{aligned} R^*(\dot\mu,\dot\alpha,\dot\beta) &= R(\mu,\alpha,\beta) \\ &= R(\mu) + R(\alpha|\mu) + R(\beta|\mu,\alpha) \\ &= 392 + 8 + 11\tfrac{7}{11} = 411\tfrac{7}{11} \end{aligned} \qquad \text{P35}$$

Because the $*$ in the symbol $R^*(\dot\mu|\dot\alpha,\dot\beta)_\Sigma$ is there to indicate that the Σ-restrictions are those involving not just $\dot\alpha$ and $\dot\beta$, but the full model involving $\dot\mu$, $\dot\alpha$ and $\dot\beta$ (only one Σ-restriction would be needed if there were no $\dot\mu$),

$$R^*(\dot\alpha,\dot\beta)_\Sigma \neq R(\alpha,\beta) \ .$$

And $R^*(\dot\alpha,\dot\beta)_\Sigma$ is calculated by (i) deleting $\dot\mu$ and the $\dot\mu$-equation from the normal equations (**) for the $\dot\mu$, $\dot\alpha$, $\dot\beta$ model, (ii) solving the equations so modified, and (iii) using the R-algorithm on the solution and the right-hand sides. This gives $R^*(\dot\alpha,\dot\beta)_\Sigma$. (This is effectively the procedure that is used in computing routines that use Σ-restrictions, such as SAS HARVEY and BMDP2V.) Carrying out (i) on (**), namely deleting $\dot\mu$ and the $\dot\mu$-equation yields

$$\begin{bmatrix} 8 & 1 & -1 \\ 1 & 5 & 2 \\ -1 & 2 & 5 \end{bmatrix} \begin{bmatrix} \tilde{\dot\alpha}_1 \\ \tilde{\dot\beta}_1 \\ \tilde{\dot\beta}_2 \end{bmatrix} = \begin{bmatrix} -8 \\ 10 \\ 4 \end{bmatrix} \quad \text{with solution} \quad \begin{bmatrix} -15/11 \\ 27/11 \\ -5/11 \end{bmatrix} .$$

The R-algorithm applied to this gives

$$R^*(\dot\alpha,\dot\beta)_\Sigma = \frac{-15}{11}(-8) + \frac{27}{11}(10) + \frac{-5}{11}(4) = 33\tfrac{7}{11}$$

and so

$$R(\dot\mu|\dot\alpha,\dot\beta) = 411\tfrac{7}{11} - 33\tfrac{7}{11} = 378 \ .$$

(v) Numerator of an F-statistic

In terms of the traditional, over-parameterized (no interaction) model

$$\dot{\mu} = \mu + \tfrac{1}{2}(\alpha_1 + \alpha_2) + \tfrac{1}{3}(\beta_1 + \beta_2 + \beta_3) \; . \qquad (29), \text{ P11}$$

For testing $H : \dot{\mu} = 0$ we equivalently consider $H : 6\dot{\mu} = 0$, or

$$H : 6\mu + 3(\alpha_1 + \alpha_2) + 2(\beta_1 + \beta_2 + \beta_3) = 0 \; .$$

The paper BU-668-M develops a form of the "indirect" calculation that is directly applicable to non-full rank models (P6-8). It is illustrated on P31-32. For purposes of that illustration the parameters μ, α_1, α_2, β_1, β_2, β_3 are used in the sequence α_1, α_2, μ, β_1, β_2, β_3, so we here consider our hypothesis rewritten as

$$H : 3(\alpha_1 + \alpha_2) + 6\mu + 2(\beta_1 + \beta_2 + \beta_3) = 0 \; .$$

A solution to the normal equations is

$$\underset{\sim}{b}^{o} = \begin{bmatrix} \alpha_1^o \\ \alpha_2^o \\ \mu^o \\ \beta_1^o \\ \beta_2^o \\ \beta_3^o \end{bmatrix} = \underbrace{\frac{1}{33}\begin{bmatrix} 18 & . & . & -12 & -6 & -9 \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ -12 & . & . & 19 & 4 & 6 \\ -6 & . & . & 4 & 13 & 3 \\ -9 & . & . & 6 & 3 & 21 \end{bmatrix}}_{\underset{\sim}{G}} \begin{bmatrix} 24 \\ 32 \\ 56 \\ 24 \\ 18 \\ 14 \end{bmatrix} = \begin{bmatrix} -2\frac{8}{11} \\ 0 \\ 0 \\ 9\frac{9}{11} \\ 6\frac{10}{11} \\ 8\frac{4}{11} \end{bmatrix} . \qquad (A7), \text{ P31}$$

Rewriting the hypothesis as $\underset{\sim}{K}'\underset{\sim}{b} = \underset{\sim}{m}$ we have $\underset{\sim}{K}' = [3 \;\; 3 \;\; 6 \;\; 2 \;\; 2 \;\; 2]$ and $\underset{\sim}{m} = \underset{\sim}{0}$ so that

$$Q = (\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})'(\underset{\sim}{K}'\underset{\sim}{G}\underset{\sim}{K})^{-1}(\underset{\sim}{K}'\underset{\sim}{b}^o - \underset{\sim}{m})$$

is

$$Q = \frac{[3(-2\frac{8}{11}) + 2(9\frac{9}{11}) + 2(6\frac{10}{11}) + 2(8\frac{4}{11})]^2}{\frac{1}{33}\{9(18) + 4(19+13+21) + 2[6(-12-6-9) + 4(4+6+3)]\}} = 378 \; .$$

With Interaction Model: Σ-restrictions

$$Q_{\dot{\mu}} = R^{*}(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} = Q \quad \text{for } H: \mu + \bar{\alpha}_{.} + \bar{\beta}_{.} + \bar{\gamma}_{..} = 0 . \qquad (29), \text{ P11}$$

Normal equations (A12) [P34 and N79] have solution

$$\hat{\underset{\sim}{b}} = \frac{1}{72}\begin{bmatrix} 10 & 0 & -1 & -1 & -3 & 3 \\ 0 & 10 & -3 & 3 & -1 & -1 \\ -1 & -3 & 19 & -8 & -3 & 0 \\ -1 & 3 & -8 & 19 & 0 & 3 \\ -3 & -1 & -3 & 0 & 19 & -8 \\ 3 & -1 & 0 & 3 & -8 & 19 \end{bmatrix}\begin{bmatrix} 56 \\ -8 \\ 10 \\ 4 \\ 18 \\ 4 \end{bmatrix} = \begin{bmatrix} 7 \\ -10/6 \\ 1 \\ -1 \\ 10/6 \\ 10/6 \end{bmatrix}, \qquad (\text{A14}), \text{ P35}$$

(ii) "Indirect"

$$Q_{\dot{\mu}} = \hat{\mu} T^{-1}_{\mu\mu} \hat{\mu} = 7(10/72)^{-1}(7) = 352.8 .$$

Note: $R(\mu) = 392$ for all models

$\neq Q_{\dot{\mu}} = 378$ for the no-interaction model

$\neq Q_{\dot{\mu}} = 352.8$ for the with-interaction model.

(iii) The $R(\cdot|\cdot)$ procedure

$$R^{*}(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} = R^{*}(\dot{\mu},\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} - R^{*}(\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} .$$

For the same kind of reasoning as on page 83

$$R^{*}(\mu,\alpha,\beta,\gamma)_{\Sigma} = R(\mu,\alpha,\beta,\gamma) = 448, \text{ from bottom of page 78.}$$

Then $R^{*}(\alpha,\beta,\gamma)_{\Sigma}$ comes from deleting $\dot{\mu}$ and the $\dot{\mu}$-equation from the normal equations (A12) [P34 and N79] to yield

$$\begin{bmatrix} 8 & 1 & -1 & 1 & 1 \\ 1 & 5 & 2 & 1 & 0 \\ -1 & 2 & 5 & 0 & -1 \\ 1 & 1 & 0 & 5 & 2 \\ 1 & 0 & -1 & 2 & 5 \end{bmatrix}\begin{bmatrix} \tilde{\dot{\alpha}}_1 \\ \tilde{\dot{\beta}}_1 \\ \tilde{\dot{\beta}}_2 \\ \tilde{\dot{\gamma}}_{11} \\ \tilde{\dot{\gamma}}_{12} \end{bmatrix} = \begin{bmatrix} -8 \\ 10 \\ 4 \\ 18 \\ 4 \end{bmatrix} \quad \text{with solution} \quad \frac{1}{30}\begin{bmatrix} -50 \\ 51 \\ -9 \\ 113 \\ -13 \end{bmatrix} .$$

The R-algorithm applied to this gives

$$R^{*}(\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} = \frac{1}{30}[-50(-8) + 51(10) - 9(4) + 113(18) - 13(4)] = 95.2$$

and so

$$R^{*}(\dot{\mu}|\dot{\alpha},\dot{\beta},\dot{\gamma})_{\Sigma} = 448 - 95.2 = 352.8 .$$

(v) Numerator of an F-statistic

The hypothesis $H: \dot{\mu} = 0$ is equivalent to

$$H: 6\mu + 3(\alpha_1 + \alpha_2) + 2(\beta_1 + \beta_2) + (\gamma_{11} + \gamma_{12} + \gamma_{13} + \gamma_{21} + \gamma_{22} + \gamma_{23}) = 0. \qquad (29), \text{ P11}$$

Hence

$$\underset{\sim}{K}' = [6 \quad 3 \quad 3 \quad 2 \quad 2 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1 \quad 1] .$$

Since $$\underset{\sim}{b}^{o'} = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 8 \quad 6 \quad 2 \quad 8 \quad 6 \quad 12] \qquad \text{P30}$$

and $$\underset{\sim}{G} = \text{diag}\{0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \tfrac{1}{2} \quad 1 \quad 1 \quad 1 \quad \tfrac{1}{2} \quad 1\} ,$$

$$Q = \frac{(8 + 6 + 2 + 8 + 6 + 12)^2}{\frac{1}{2} + 1 + 1 + 1 + \frac{1}{2} + 1} = \frac{42^2}{5} = 352.8 .$$

Data With Empty Cells

Methods of calculation are essentially the same.

One difference is that Q for hypotheses tests involves hypotheses that are not as mathematically symmetric as with data wherein all cells are filled.

Example

2,4,6	4,6	5	27
12,8	11,7	-	38
32	28	5	65

re-sequenced

	β_1	β_3	β_2	
α_2	12,8	-	11,7	38
α_1	2,4,6	5	4,6	27
	32	5	28	65

"Indirect"

The resequencing of the data is as required by SAS HARVEY, so that the last row and column have no empty cells. (If no such row and column exist, and the model has interactions, the output is an error message.) Then using $\dot{\alpha}_1 = -\dot{\alpha}_2$ and $\dot{\beta}_2 = -\dot{\beta}_1 - \dot{\beta}_3$, the normal equations for the with-interaction, Σ-restricted, model are

$$\begin{bmatrix} 10 & -2 & 1 & -3 & -1 \\ -2 & 10 & -1 & -1 & 1 \\ 1 & -1 & 9 & 4 & -1 \\ -3 & -1 & 4 & 5 & 0 \\ -1 & 1 & -1 & 0 & 9 \end{bmatrix} \begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_2 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_3 \\ \hat{\dot{\gamma}}_{21} \end{bmatrix} = \begin{bmatrix} 65 \\ 11 \\ 4 \\ -23 \\ 0 \end{bmatrix},$$

with solution

$$\begin{bmatrix} \hat{\dot{\mu}} \\ \hat{\dot{\alpha}}_2 \\ \hat{\dot{\beta}}_1 \\ \hat{\dot{\beta}}_3 \\ \hat{\dot{\gamma}}_{21} \end{bmatrix} = \frac{1}{864} \begin{bmatrix} 155 & 39 & -89 & 172 & 3 \\ 39 & 99 & -21 & 60 & -9 \\ -89 & -21 & 203 & -220 & 15 \\ 172 & 60 & -220 & 464 & -12 \\ 3 & -9 & 15 & -12 & 99 \end{bmatrix} \begin{bmatrix} 65 \\ 11 \\ 4 \\ -23 \\ 0 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 43 \\ 15 \\ -1 \\ 2 \\ 3 \end{bmatrix}.$$

and so $$Q_{\dot{\mu}} = (43/6)^2(155/864)^{-1} = \frac{44376}{155} = 286\frac{46}{155} = 286.29677 .$$

Numerator of an F-statistic

A solution vector of the normal equations for the over-parameterized model with parameter vector

$$\underset{\sim}{b}' = [\mu \quad \alpha_1 \quad \alpha_2 \quad \beta_1 \quad \beta_2 \quad \beta_3 \quad \gamma_{11} \quad \gamma_{12} \quad \gamma_{13} \quad \gamma_{21} \quad \gamma_{22}]$$

is

$$\underset{\sim}{b}^{o'} = [0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 4 \quad 5 \quad 5 \quad 10 \quad 9\ ]$$

with

$$\underset{\sim}{G} = \text{diag}\{0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad \tfrac{1}{3} \quad \tfrac{1}{2} \quad 1 \quad \tfrac{1}{2} \quad \tfrac{1}{2}\ ] .$$

The hypothesis $H : \dot{\mu} = 0$ corresponds to

$$H : \mu + \frac{1}{2}(\alpha_1 + \alpha_2) + \frac{1}{3}(\beta_1 + \beta_2 + \beta_3) + \frac{1}{12}(\gamma_{11} + \gamma_{12}) + \frac{1}{3}\gamma_{13} + \frac{1}{4}(\gamma_{21} + \gamma_{22}) = 0$$

so that

$$\underset{\sim}{K}' = \left[1 \quad \frac{1}{2} \quad \frac{1}{2} \quad \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{3} \quad \frac{1}{12} \quad \frac{1}{12} \quad \frac{1}{3} \quad \frac{1}{4} \quad \frac{1}{4}\right] .$$

Hence Q is

$$Q = \frac{[\frac{1}{12}(4 + 5) + \frac{1}{3}(5) + \frac{1}{4}(10 + 9)]^2}{\frac{1}{144}\left(\frac{1}{3} + \frac{1}{2}\right) + \frac{1}{9}(1) + \frac{1}{16}\left(\frac{1}{2} + \frac{1}{2}\right)} = \frac{44376}{155} = 286\,\frac{46}{155} .$$

In looking further at details of this example (e.g., p. 81), or indeed at those of other small examples, one should not be led into false comfort that interpretation of parameters in a Σ-restricted model is always easy or useful. It is both of these things in the case of balanced data and it is mostly that way for the case of all cells filled; but for the case of some cells empty it is generally neither easy nor useful. Furthermore, the exact interpretation depends upon just which cells are empty. The example considered is relatively simple, in terms of the pattern of empty cells, and yet interpretation of the Σ-restricted model is not easy. For more extensive data, interpretation is even more complicated.

8.3. Other restrictions

a. W-restrictions

Constraints sometimes useful for solving normal equations in the unrestricted model are those which involve weighted sums such as $\Sigma n_{i.}\alpha_i^o$. Analogous restrictions on the model are called the W-restrictions. An example is

$$\sum_{i=1}^{a} n_{i.}\alpha_i = 0\ , \qquad \sum_{i=1}^{a} n_{ij}(\alpha_i + \gamma_{ij}) = 0 \quad \forall j$$

$$\sum_{j=1}^{b} n_{.j}\beta_j = 0\ , \qquad \sum_{j=1}^{b} n_{ij}(\beta_j + \gamma_{ij}) = 0 \quad \forall i$$

used by Speed et al. [J.A.S.A., 1978]. They indicate that relationships to the classical analysis of variance are of the form

$$R^{*}(\alpha|\mu,\beta,\gamma)_W = R(\alpha|\mu) \quad \text{for the all-cells-filled case.}$$

Restrictions of this form are used in SPSS.

b. <u>0-restrictions</u>

Other restrictions sometimes employed are those which put some parameters to zero, analogous to a procedure for solving normal equations (e.g., LM 213). They can be called the 0-restrictions. An example of these considered by Speed and Hocking [Amstat, 1976], which we call the 0_{11}-restrictions, is

$$\alpha_1 = 0 \quad \text{and} \quad \gamma_{1j} = 0 \ \forall j$$

$$\beta_1 = 0 \quad \text{and} \quad \gamma_{i1} = 0 \ \forall i .$$

A generalization, to be called the 0_{kt}-restrictions, is

$$\alpha_k = 0 \quad \text{and} \quad \gamma_{kj} = 0 \ \forall j$$

$$\beta_t = 0 \quad \text{and} \quad \gamma_{it} = 0 \ \forall i .$$

Then for the all-cells-filled case

$R^*(\alpha|\mu,\beta,\gamma)_{0_{kt}}$ can be used to test $H: \beta_j + \gamma_{kj}$ all equal.

In view of the restrictions, this hypothesis is $H: \beta_j = 0$, for all j; but it is, of course, a hypothesis of equality of columns tested over a specified row, namely the k'th row.

APPENDIX

Papers of the Biometrics Unit, Cornell University, which supplement these notes.

BU-451-M: Hypothesis testing in restricted linear models: correcting an error. (S. R. Searle, 1973, revised 1978.)

BU-501-M: Testing non-testable hypotheses in linear models. (S. R. Searle, 1973.)

BU-533-M: Restrictions on models and constraints on solutions in analysis of variance. (S. R. Searle, 1974.)

BU-342: Alternative covariance models for the 2-way crossed classification. (S. R. Searle, 1978.) Communications in Statistics A8, 799-818, 1979.

BU-343: Relationships between the estimable functions of SAS GLM output for unbalanced data and the hypotheses tested by traditional-style F-statistics. Proceedings 4th Annual SAS Users' Group International Conference, 196-208, 1979.

BU-668-M: Some computational and model equivalencies in analysis of variance of unequal-subclass-numbers data. (S. R. Searle, H. M. Speed and H. V. Henderson, 1979.)

BU-672-M: Expected marginal means in the linear model. (S. R. Searle, G. A. Milliken and F. M. Speed, 1979.)

BU-682-M: On Hadamard products, interactions, covariates and computer routines for linear models. (S. R. Searle and H. V. Henderson, 1979.)